Applications of Grid-Based Likelihood Methods to CMB Analysis

Kristin Mikkelsen

Institute of Theoretical Astrophysics
University of Oslo

2014

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1446

ISSN 1501-7710

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika Publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

Preface

Of all the possible career paths out there how does one end up as a cosmologist? In my case this is a result of having a mathematically inclined brain that deals better with a fact-based world like cosmology, than with the philosophical world, such as art or literature where your imagination is the only restriction. The field of cosmology certainly fulfills this fact-driven requirement, but is also appealing as a field that constantly finds itself at the brink of some epic discovery to deepen our understanding of the Universe. Perhaps the question should instead be why more people don't choose to get involved with this field.

Studying the properties of the Universe through data analysis is not the easiest task, but delving into this scene is definitely an interesting and endless adventure with a steep learning curve that for me has been filled with confusion, epiphanies, head scratching and moments of bliss. For non-cosmologists the endeavor may seem like a mixture of magic and great wisdom, but of course cosmology is just like any other field, merely on a (really) large scale. Cosmological data analysis is primarily the application of statistical techniques to data, in the same manner as that for other data, like population statistics, global warming or medical studies. The difficulty arises in figuring out how the Universe functions; what does it contain, what types of physics does it obey and how to interpret the results, as well as only having one single Universe to study.

The red thread through my Ph.D-thesis is the analysis of CMB data based on maximum likelihood search and in particular the application of grid-based sampling methods. The fields include cosmological parameter estimation, likelihood evaluation techniques and statistical analyses, but the work has also encompassed data processing based on a slightly different statistical framework. Furthermore, in the spirit of the traditional astronomy I have also partaken in spectroscopic observations of galaxy clusters at two different telescopes.

Acknowledgements

The most obvious person to thank is my supervisor Hans Kristian Eriksen, who's insight and skill at finding interesting and challenging projects are invaluable, and for always taking time to explain/discuss problems, results and ideas. Without his guidance this would never have been possible. The same goes for my former colleague Sigurd Næss who was an important collaborator, brainstorming partner and technical support, and without him and Frode Hansen my teaching duties and the endless assignment corrections would not have been quite as enjoyable.

A number of subsets of colleagues at the institute also deserve a thank you for having contributed to my academic life in one way or another, and the first of these are my former office mates Nicolaas Groeneboom and Øystein Rudjord, who in many ways were the welcoming commitee to ITA - showing me the ropes, and who formed the basis of my social life in Oslo. I would also like to thank Håkon Dahle for sending me to the La Palma and La Silla observing sites to perform observations. I am also grateful to my current officemates, Philip Bull, Unni Fuskeland, Eirik Gjerløw, Tone Ruud and Dag Sverre Seljeboth for helpful discussions and for a great work environment.

I also want to thank my housemates during the last four years; Patrick Antolin, Tuva Hope, Astrid Momyr and Lillian Smestad for the feeling of "home", and Nicolaas, Øystein and Zofia Czyczuła Rudjord for this summer's sailing adventure, as well as Kosovare Olluri and Ingeborg Ligaarden. To all of you I'd like to say thanks for the laughs, home-cooking get-togethers, distractions, random activities, work sessions and means of procrastination.

I am also grateful to my family, and in particular to my parents, Karin Eriksen and Karl Mikkelsen, for letting me make my own decisions and mistakes, and for providing a gentle nudge or proper shove every so often. My brother Bjørn Mikkelsen deserves a spesial mention too, for inspiring me to learn and be better from an early age, and for showing me it's okay to follow my own path, even if it's slightly crazy.

The cosmology groups at the University of Oxford (UK) and Jet Propulsion Laboratory (USA) where I have spent a considerable amount of time also deserve a mention. Lastly, a special thanks to the Norwegian Programming Center for providing me with a quiet work space to write this monster, and to Bjørn and Nicolaas for diligent proof-reading.

Contents

Part I

Cosmological Framework

Chapter 1

Introduction to Cosmology

Mankind's curiosity and craving for knowledge has lead to the development and deepened understanding of a great number of ideas and fields, not to mention the wealth of technological advances that are part of our modern way of life. The subfield of astrophysics that is concerned with the study of the largest scales imaginable, namely the entire Universe, is called cosmology and includes investigation of how the Universe came to be, its evolution both in the past and into the future, and the different constituents it contains.

One can easily imagine that the first generations of astronomers looked up at the sky and marveled at the wonders they saw, similar to the romantic idea that all modern astronomers spent their childhood gazing at the stars, wondering what they are and how they were created. The knowledge of the Universe in the childhood of astronomy was greatly limited and in many cases rather flawed, and one can therefore not help but admire the courage and resolve of these scientists as they delved into the unknown; going where no scientists had gone before by trying to reach the stars. Although our view of the Universe has changed significantly since the time of those early pioneers the basic questions have remained the same: why are we here, where and when is here, how did here come to exist, what are we made of and what is all the other stuff made of. My prediction is that we will never reach the end of the quest for answers as it seems that for every question solved several more pop up to take its place, but mankind's need to know will drive the field of cosmology, and the rest of astrophysics, for many years to come. Hopefully the next few decades will be as rewarding as the last couple have been, and maybe, just maybe, the race for knowledge is almost over.

In the history of astronomy the "truth" has changed drastically a number of times, for example back in the 15th century the solar system was believed to be a geocentric system with the Earth at the center and all other bodies orbited around it. However, after detailed observations and laborious calculations, the idea of a heliocentric model[1] was put forward and eventually

[1]The heliocentric model is also known as the Copernican model.

accepted, which places the sun at the center of the solar system with Earth and the other planets orbiting it. The truth was rewritten. Interestingly, moving the Earth away from the center of the solar system was the first step towards the Copernican principle, which states that our place in the Universe is not a special one, a concept that is of great importance in modern cosmology. Another drastic modification in astronomy came in the 1920s and involves our view of the extent of the Universe. Roughly a hundred years ago the Universe was believed to consist of only the Milky Way galaxy, putting a rather tight constraint on the cosmological models in the astrophysical community. The discovery of nebulae and galaxies outside our own lead to the conclusion that we live in a universe of unimaginable size, and left astronomers pondering the relevance our existence.

So where do we stand today? What do the experts say about the Universe that we are such a tiny part of? Before reviewing the currently adopted cosmological model let us rewind 50 years to a time when the battle between two very different cosmological theories was going strong, the victor of which, with some slight modifications, has survived to present date. In the 60s it was known that the Universe was expanding based on luminosity measurements indicating that galaxies were moving away from us[2]. However, the two main classes of models to describe this expanding Universe were exceptionally different. In brief terms, there was the Steady State theory [3, 4] describing an expanding non-evolving Universe where matter is continually created, a universe with no beginning and no end; a universe that would remain steady for all of time. The other contender was the Big Bang theory [5, 6, 7] with an evolving Universe that started off being very hot and dense [8, 9], but cooled as the Universe expanded, causing an evolution of its constituents. This could lead to several radically different futures depending on the total weight of the Universe. Both models are dependent on the creation of matter in some way; the Steady State theory relies on a continuous creation of matter for all of time to keep the local matter content constant, and the Big Bang theory requires that at the start of the Universe all of the matter in the Universe was created.

The observational evidence at the time was not consistent and seemed to favor different models depending on the observations. Hubble's expansion measurements from 1929 [2] lead to the conclusion that the time since the Big Bang was smaller than the age of the solar system, which was clearly wrong . This therefore favoured the Steady-State theory which relates Hubble expansion to continuous matter formation. On the other hand, cluster counts at large distances in the early 1960s were in disagreement with those predicted by the Steady-State theory [10], and so were the abundance of light elements [8, 9]. Eventually the controversy surrounding the low age calculated from the expansion measurements were remedied when it was

[2]Observations courtesy of Henrietta Leavitt [1] and analysis by Edwin Hubble [2].

discovered that the observational sources were found to be slightly different objects than originally thought, thereby changing the luminosity relationship and in turn increasing the age of the Universe. Over time the observational evidence, like the age of the Universe and cluster counts, were found to be more in favour of Big Bang theory and disfavoring the Steady-State theory.

However, it was not until 1964 that the decisive blow came to properly tip the scale in favor of the Big Bang theory. This was the serendipitous discovery of the cosmic microwave background (CMB) radiation by Arno Penzias and Robert Woodrow Wilson while working with a radio receiver at Bells Laboratories in New Jersey. They found a persistent radiation signal of around 3 K across the whole sky, which they could not explain after accounting for all the contributions from their equipment, excluding contamination from Earth and the solar system and even relocating some nesting pigeons. At the same time, about an hour drive away at Princeton University a team lead by Robert H. Dicke, Jim Peebles, and David Wilkinson were preparing an experiment to search for a relic signal from the Big Bang. If the Big Bang theory help any merit they theorized that this relic signal should have a temperature of a few Kelvin and therefore be observable in the radio regime. Penzias was made aware of the theory explaining the puzzling observations and after some communication the two groups decided to published the theory of a relic CMB signal [11] and the observational evidence with a temperature of 3.5 ± 1.0 K, [12] in the same journal. However, only Penzias and Wilson received the Nobel Prize[3] for the discovery of this background signal. Of the two contenders only the Big Bang theory predicted the existence of this relic radiation, which left the Steady State theory found wanting and the Big Bang theory has reigned every since.

[3]Penzias and Wilson received the 1978 Nobel Prize for the discovery of the cosmic microwave background radiation, which they shared with Pyotr Kapitsa for unrelated work on low-temperature physics.

Chapter 2

Big Bang Theory - a Brief Journey Through Time

The original Big Bang theory was independently proposed by Alexander Friedmann [5, 6] (in German) and Georges Lemaître [7] (in French) in the late 1920s, and given the name "Big Bang" by the Steady State theory advocate Fred Hoyle[1] when describing the early highly expansive phase of the theory during a BBC radio show[2] in 1949. The catchy term stuck and is still in use, even though it often leads to misunderstandings amongst the general public with regards to the nature of this "explosion", and particularly where in space it supposedly took place - the answer to which, rather frustratingly, can be both everywhere and nowhere, depending on how you look at it.

The details and explanations in this thesis are mainly taken from books by Dodelson [13] and Peacock [14], and introductory papers to cosmology by Gumjudpai [15], Reid et al. [16] and Brandenberger [17].

2.1 Expansion of the Universe

As mentioned in chapter 1, astronomical observations indicate that almost all galaxies are moving away from us and each other. In other words, the distance (d) between most[3] pairs of galaxies is increasing and the Universe as a whole is expanding. The relative expansion of the Universe at time t can be quantized by the scale factor $a(t)$, which is set to 1 at an appropriate reference time t_0. The commonly adopted notation of using a subscripted zero for any quantity at this reference time will be followed, for example the distance at the reference time $d(t_0)$ and the corresponding scale factor

[1] Fred Hoyle died in 2001 and never came to accept the Big Bang theory.

[2] Reprinted in April 1949 in The Listener, a BBC magazine.

[3] Some galaxies are in such close proximity that their mutual gravitational attraction locks them together. An example of such a bound system is the Milky Way galaxy and the neighboring spiral galaxy Andromeda, which will eventually collide.

$a(t_0)$ are represented by d_0 and a_0, respectively. The distance between two galaxies at time t is related to the distance between them at the reference time through

$$\frac{d(t)}{a(t)} = \frac{d_0}{a_0} = d_0, \tag{2.1}$$

where we use that $a_0 = 1$. For convenience the reference time t_0 is chosen to be present time, thus $a(t)$ is the factor by which a distance in the Universe at time t is greater or smaller than that today. The physical distance at time t can therefore be written in terms of a comoving distance, x, and the scale factor as

$$d(t) = a(t)x, \tag{2.2}$$

where the comoving distance is part of the comoving coordinate system which is defined to follow the expansion of the Universe in a manner such that the expansion can be interpreted as a stretching of space itself. The physical distance increases with expansion, in the same manner as the scale factor, but the comoving distance remains constant. This is illustrated in figure 2.1 where the blue and red points are two galaxies moving apart as a result of the expansion and the solid and dotted grids correspond to the comoving and physical grids, respectively. The leftmost figure shows overlapping grids at a time when the scale factor of the Universe equaled 1, in which case the physical separation between the two galaxies was identical to the comoving separation. The figure on the right shows the two grids at time t_2 when the scale factor has increased to 1.5, which means that the physical distance has increased by the same factor. However, the comoving grid follows the expansion and therefore the comoving separation has the same value as for the Universe with $a(t_1) = 1$.

In an astrophysical context it is also useful to introduce the concept of redshift, which describes the stretching of an electromagnetic wavelength due to the expansion of the Universe. Redshift, like the scale factor, is an evolving quantity and relates the wavelength at emission, λ_{em}, to that at observation, λ_{ob}, through the relationship

$$z = \frac{\lambda_{\text{ob}}}{\lambda_{\text{em}}} - 1, \tag{2.3}$$

which gives the redshift of the electromagnetic source at time t_{em} and can be thought of as a representations of distance to this source. The expansion of space and redshift are related through

$$z = \frac{a_{\text{ob}}}{a_{\text{em}}} - 1, \tag{2.4}$$

thus the redshift increases with decreasing scale factor at emission.

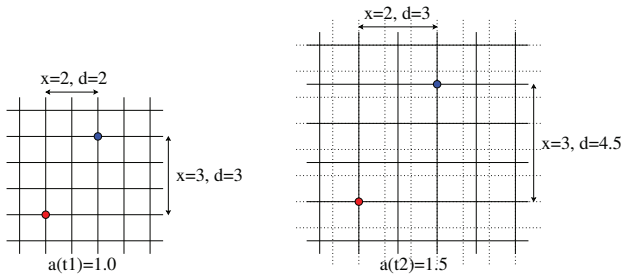


Figure 2.1: Schematic showing the expansion of the comoving grid (solid grid lines) as a result of the expansion of space, and it's relation to the static physical grid (dotted grid lines), for two-dimensional space containing two galaxies (red and blue dots). Left: Grids and galaxies at the reference time, t_1, where the physical and comoving grids overlap, and the separations are equal. Right: Grids at time t_2, where the physical separation between the galaxies has increased by a factor $a(t_2) = 1.5$, whereas the comoving separation has remained constant, since the coordinate grid has expanded.

Furthermore, not only did the analysis of Henrietta Leavitt's stellar brightness catalogue [1] by Edwin Hubble in the late 1920s result in evidence for an expanding Universe [2], but also one in which more distant galaxies are receding faster than those closer by. This may seem obvious, but only if one follows the Copernican principle which, as mentioned, can be boiled down to the statement that to our place in the Universe is not a special one and therefore that observations from any point in space should result in equally valid descriptions of the Universe. Thus, seen from any point in the Universe, all galaxies are moving away, but more distant galaxies are doing so at a higher velocity in order to account for the mutual recession speed between any two galaxies. This concept is illustrated in figure 2.2 where the central figure shows a set of galaxies in the Universe at the reference time t_1, which is underlayed in the figures on either side. The figures on the left and right show the positions and velocities of the galaxies at time t_2 from the point of view of observers in the blue and red galaxy, respectively. Their own galaxy appears to them to be at rest and the galaxy of the other observer moves away at speed v. Additionally, the remaining galaxies recede at different velocities depending on their distance from the observer in a manner such that a galaxy twice as distant moves twice as fast and one three times as far away moves thrice as fast. This lead to the famous Hubble's Law relating velocity and distance through

$$\boldsymbol{v} = Hd, \tag{2.5}$$

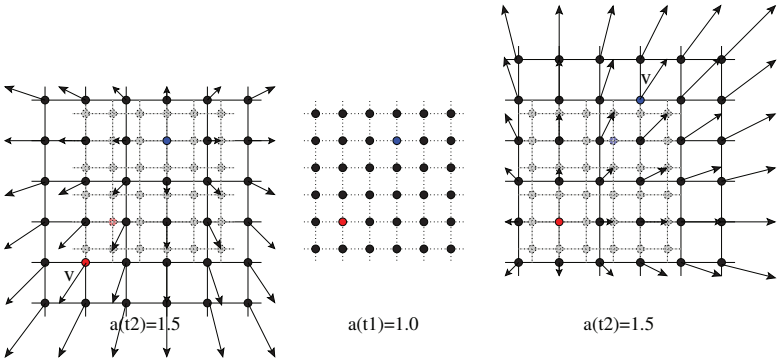


Figure 2.2: Illustration of Hubble's law, which states that the recession velocity of an object is directly proportional to its separation distance. The central figure shows the physical positions of a set of galaxies when $a(t_1) = 1$, which is underlayed for the illustrations on either side. The left and right figures show the positions and velocities of the galaxies when $a(t_2) = 1.5$ from the point of view of the blue and red galaxy, respectively.

where $\boldsymbol{v}$ and d are the velocity of and distance to a galaxy, and H, known as the Hubble constant, represents the expansion rate of the Universe at the time of observation. The Hubble constant is only a constant in space[4] and not in time, and is therefore more accurately described as an evolving quantity, dubbed the Hubble rate, which relates to the scale factor through

$$H(t) = \frac{da(t)/dt}{a(t)} = \frac{\dot{a}(t)}{a(t)}, \tag{2.6}$$

where the superscripted dot represents derivative with respect to time, a notation that will be followed throughout this thesis. As can be seen from equation (2.6) the Hubble rate is directly related to the scale factor and its time derivative, thereby quantifying how fast the Universe is expanding with respect to its current size, at a given point in time. The most recent CMB measurements with the *WMAP* and *Planck* satellites gave Hubble rates today of $H_0 = 70.0 \pm 2.2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ [18] and $H_0 = 67.3 \pm 1.2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ [19], respectively.

Based on galaxy measurements it is clear that a universe following a Big Bang model expands as it ages, and therefore cools to conserve energy - much like the temperature of gas in a box would decrease as the box expands, to compensate for the increase in volume. This leads to an evolution in the scale factor, temperature, energy density etc. of the Universe, which is dependent

[4]The Hubble rate is only a constant in space if all spatial directions are assumed to behave in the same manner, otherwise it can vary in spatial directions as well: $H(t, x, y, z)$.

on the dominant component in the Universe at a given time. The dominant component also changes with time leading to transitions in the Universe's evolution, which we will come back to in section 2.4.

2.2 Einstein's theory of General Relativity

The Big Bang theory leans heavily on Einstein's theory of general relativity (GR) which unifies time and three-dimensional space into four-dimensional spacetime, and sets an upper limit for the flow of information at the speed of light in vacuum, c. This limitation means that the transfer of any kind of information (force, particle, object, energy, etc.) requires time to travel through space, and the resulting time lag is dependent on the distance traveled. The theory is named after its creator Albert Einstein, who published his revolutionary theory in 1915 [20, 21], and was used to describe the birth and evolution of the Universe and to predict its fate. Within the framework of GR the nature of gravitation is not explained as a result of the pull of a gravitational field set up by massive objects, but is rather described as a geometric property resulting from massive objects curving spacetime, into which other objects fall as they follow geodesics[5]. This is illustrated for two-dimensional space in figure 2.3 for a light ray passing close to the Sun on its way to Earth. The energy density of the Sun curves space, here represented by a third spatial dimension, thus the light reaching Earth from point A follows a straight geodesic along the curved surface, which would look like a curved path in the two-dimensional space. The light reaching Earth appears to originate from point B when GR is not considered, which relates to the experiment lead by Arthur Eddington that was performed in 1919 to test and ultimately prove the theory of general relativity during a solar eclipse [22][6].

2.2.1 Einstein's field equation

In GR gravitation is described through Einstein's field equations which are a set of 10 coupled tensor equations[7] that describe how the interaction of gravity arises as a consequence of energy curving spacetime, written as

$$G_{\mu\nu} - g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4} T_{\mu\nu}, \qquad \text{with} \tag{2.7}$$

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \tag{2.8}$$

[5] A geodesic is the straightest path in curved spacetime.

[6] Eddington and his team took images of the position of stars during a total solar eclipse and compared these to images without the Sun present. A shift in stellar position was observed indicating that the gravitational potential of the Sun affects the path of light rays.

[7] By using the 4 Bianchi identities the number of coupled equations is reduced to 6.

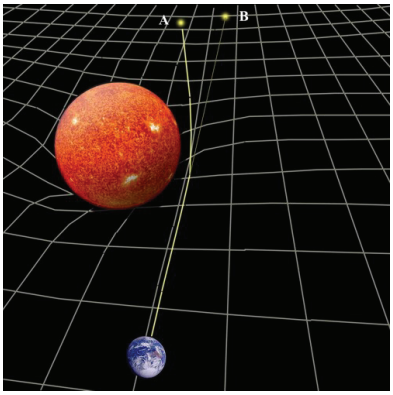


Figure 2.3: Two-dimensional space curves into a third dimension in response to the energy density of the Sun, thereby causing light received at Earth from the star at A to follow a curved path. The source therefore appears to be located at position B. Source: Time Travel Research Center.

where $G_{\mu\nu}$, $g_{\mu\nu}$, $T_{\mu\nu}$ and $R_{\mu\nu}$ are, respectively, the four-by-four dimensional Einstein, metric, energy-momentum and Ricci curvature tensors. Λ, G, c and R are the cosmological constant, Newton's gravitational constant, speed of light in vacuum and Ricci scalar curvature, respectively. The metric tensor is used to compute physical distances from coordinates and is directly related to both the Ricci curvature tensor and scalar curvature, that all vary for different kinds of parameter spaces. As shown in equation (2.8) these three combine to give the Einstein tensor, describing the overall geometry of the Universe. The energy-momentum tensor, rather obviously, describes the energy and momentum of the constituents of the Universe. The left hand side of equation (2.7) represents the shape and curvature of spacetime and the right hand side the energy content of the Universe, thereby describing a relationship between energy and spacetime, and how these affect one another.

The cosmological constant, Λ, that makes an appearance in Einstein's field equation will be discussed in detail in section 2.6, and for now it will suffice to say it represents a geometric consequence of a non-constant expansion of the Universe. Einstein introduced this mysterious Λ on the energy side of equation (2.7) as a means of making his equations static, since at the time the known Universe only contained the Milky Way galaxy which was believed to remain unchanged. Without this extra term the Universe would collapse due to the gravitational attraction of matter, contrary to the observations at the time. However, adding the Λ term makes the solution unstable and it was abandoned (for a while).

2.2.2 Metrics and line elements

In two-dimensional Euclidean space the distance between two points, dl, separated by dx and dy in Cartesian coordinates, is given by

$$dl^2 = dx^2 + dy^2. \tag{2.9}$$

This distance is an invariant quantity meaning that it is independent of coordinate system and therefore always gives the same separation distance regardless of the coordinate system the observer is operating in. For example, the Cartesian coordinate system (x, y) above can be converted to the equivalent polar coordinate system (r, θ) by using $x = r\cos\theta$ and $y = r\sin\theta$, which gives

$$dl^2 = dr^2 + r^2 d\theta^2, \tag{2.10}$$

where we have used that $dx = \cos\theta dr - r\sin\theta d\theta$ and $dy = \sin\theta dr + r\cos\theta d\theta$. The metric is a useful quantity to introduce, and enables us to remove the coordinate-dependence from invariant quantities, for example through,

$$dl^2 = \sum_{i,j=1}^{2} g_{ij} dx^i dx^j = g_{ij} dx^i dx^j, \tag{2.11}$$

where the metrics are given by

$$g_{ij} = \mathrm{diag}(1,1), \quad x^1 = x \quad \text{and} \quad x^2 = y \quad \text{or} \tag{2.12}$$

$$g_{ij} = \mathrm{diag}(1,r), \quad x^1 = r \quad \text{and} \quad x^2 = \theta, \tag{2.13}$$

for Cartesian or polar coordinates, respectively. Einstein's notation has been introduced to avoid writing the summations explicitly every time, where summation is represented by repeated indices.

In four-dimensional spacetime the invariant quantity is given by the line element,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \tag{2.14}$$

where ds is called the proper time, and the Greek letters μ and ν are used to represent that four-dimensional spacetime is being considered, whereas i and j are reserved for one to three-dimensional space with no time component. The time component is not treated equivalently to the three spatial dimensions and does not add to the spatial dimensions in a straightforward manner to give the line element, but rather through

$$ds^2 = c^2 dt^2 - (dx^2 + dy^2 + dz^2) = c^2 dt^2 - dl^2, \tag{2.15}$$

when considering a static flat Universe in Cartesian coordinates, with the accompanying Minkowski metric given by $g_{\mu\nu} = \mathrm{diag}(c^2, -1, -1, -1)$. In

three-dimensional space the path a test particle takes between two points is the straight line that minimizes dl. However, in four-dimensional spacetime the path followed is that of maximal aging, meaning the path that maximizes the time measured on the particle's local watch, which is called the Principle of Maximal/Extremal Aging [23]. The line element is invariant meaning that all observers measure the same ds between two event, but the spatial and time differences between events vary from frame to frame. For example, two events for a muon (e.g. entering the atmosphere and decaying) will in the lab frame measure a time difference dt_{lab} and a spatial difference dl_{lab}. In the frame attached to the particle there is no movement, thus $dl_\mu = 0$, and the time difference between the two events is dt_μ. The line element must be conserved thus

$$c^2 dt_\mu^2 = ds^2 = c^2 dt_{\text{lab}}^2 - dl_{\text{lab}}^2, \tag{2.16}$$

and dt_μ becomes the proper time, that is often annotated by $d\tau$. The principle of maximal aging means that this quantity is maximized for test particles moving through spacetime.

2.3 Geometry of the Universe

In the previous section (2.2) gravitation was described though GR as massive objects causing spacetime to curve, however the overall geometry of the Universe can also have a global curvature of space, which is represented by k. The geometry of the Universe can be described with respect to its critical energy density, $\rho_{\text{crit}} = 3H^2/8\pi G$, through

$$\Omega_{\text{tot}} = \frac{\rho_{\text{tot}}}{\rho_{\text{crit}}}, \tag{2.17}$$

where ρ_{tot} is the Universe's total energy density. A universe with $\Omega_{\text{tot}} = 1$ is said to be flat with a curvature given by $k = 0$, although there is of course curvature on local level due to massive objects. If the total energy density of the Universe is less than or greater than the critical energy density ($\Omega_{\text{tot}} < 1$ or $\Omega_{\text{tot}} > 1$) it is said to have, respectively, an open or closed geometry with $k = -1$ or $k = +1$.

The top row of figure 2.4 shows an illustration of the three different types of global geometry, represented by a ball, sheet and saddle for the closed, flat and open geometries, respectively, together with how geometry affects the sum of angles in a triangle on the surface of these geometries ($> 180°$, $= 180°$ and $< 180°$, respectively). In the absence of local curvature the geometries affect the paths of two initially parallel light beams in different manners where for a flat universe the light rays will stay parallel, but in closed or open Universes they will converge or diverge, respectively. The geometry of the Universe will therefore also affect observations and in the case of the

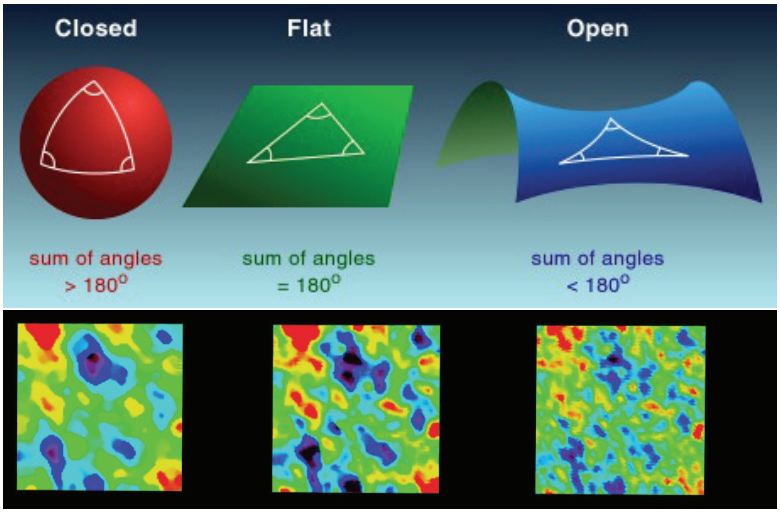


Figure 2.4: Top panels: Space is represented by a ball, sheet and saddle for the closed, flat and open geometries, respectively. The sum of the angles of a triangle changes depending on geometry as $> 180°$, $= 180°$ and $< 180°$, respectively. Bottom panels: The size of structures in the CMB radiation depends on how geometry affects parallel light rays such that structures appear to be larger for a closed Universe, unchanged for a flat Universe and smaller for an open Universe, with respect to reality. Credit: Department of Physics, The University of Hong Kong (top) and NASA / *WMAP* Science Team (bottom).

CMB radiation the geometry affects how the radiation field looks when it reaches us, as can be seen in the bottom panel of figure 2.4. In a universe with a closed geometry the CMB photons traveling through space converge due to the positive curvature and structure on the CMB sky therefore appears to be larger in size than in reality. The reverse is true for a negative curvature which causes structure to look smaller, and flat geometry obviously does not affect the structure size. Observations indicate that the Universe is very close to flat [18, 19], which, as we will see in section 2.6, causes some problems for the Big Bang model.

2.4 The FLRW Universe

Of the many assumptions made in modern cosmology, the most important is that the laws of physics are the same throughout the Universe and in time. This is described through the Cosmological Principle which states that the Universe is isotropic and homogeneous, meaning that the distribution of matter in the Universe, and therefore of physical processes, is the same in every direction and at every point in space, but not in time[8]. As a consequence of isotropy the scale factor and Hubble rate, quantifying expansion and expansion rate, should be the same in all three spatial dimensions. On the other hand, if isotropy is broken the scale factor and Hubble rate are not space invariant and must include space variation, for example the scale factor could be represented by $a(t, x, y, z)$ in Cartesian coordinates. The assumptions of the Cosmological Principle have been carefully tested and may one day result in a revolution in the cosmological community.

Under the assumption of isotropy and homogeneity an exact solution to Einstein's field equation for a contracting or expanding Universe is found using the Friedmann-Lemaître-Robertson-Walker[9] (FLRW) [5, 7, 24, 25] metric, which results in an invariant line element given by

$$ds^2 = c^2dt^2 - a(t)^2\left[\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right], \tag{2.18}$$

where k is the spatial curvature of the Universe that takes the values -1, 0 and +1 for an open, flat and close Universe, respectively. Furthermore, the spatial component of the metric in an FLRW Universe is limited to being dependent only on time, as consequence of the Cosmological principle.

Before solving the field equations, a few more pieces of the puzzle are required, namely the distribution of energy and momentum and the equations of state for the fluids in the Universe, describing the nature of a fluid through the relationship between its density, ρ, and pressure, p. For simplicity the fluids in the Universe are modeled as perfect fluids, thereby having zero viscosity and no heat conduction in comoving coordinates. Furthermore, by being perfect fluids, they have an energy momentum-tensor and equation of state given by,

$$T^{\mu}_{\nu} = \mathrm{diag}(\rho, -p, -p, -p) \qquad \text{and} \tag{2.19}$$

$$p = w\rho c^2, \tag{2.20}$$

respectively, where w is known as the equation of state parameter that takes a specific value for different fluid types. In particular for matter[10] $w = 0$,

[8] The Perfect Cosmological Principle is a generalization of the Cosmological Principle to also include time invariance, and was the cornerstone of the Steady State theory.

[9] The FLRW metric is known by many names, either by all four names combined like here, or through a variety of combinations using just two or three of these names.

[10] Matter and dust are both terms used for the total matter in the Universe, which includes both baryonic matter and dark matter.

for radiation $w = 1/3$ and for a cosmological constant $w < 0$ (negative pressure). These three fluids are the most common constituents in the Universe and represent non-relativistic matter, relativistic matter and dark energy, respectively. The latter of these is directly related to the Λ parameter introduced in Einstein's field equation, and is, as we will see, of great importance in the Universe today.

For a perfect fluid filled FLRW Universe the field equations can be reduced to a set of equations giving the evolution of the scale factor as a consequence of the constituents through

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \qquad \text{and} \tag{2.21}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\rho\left(1 + \frac{3w}{c^2}\right) + \frac{\Lambda c^2}{3}, \tag{2.22}$$

which are called the Friedmann and acceleration equations, respectively. In addition we can write the evolution of density as

$$\dot{\rho} + 3H\rho\,(1 + w) = 0, \tag{2.23}$$

which is known as the continuity equation. These three equations[11] are not independent of each other as two of them can be combined to derive the third.

The simplification of setting $c = 1$ to remove the speed of light from the equations is used for the rest of this chapter, which means that distance and time are equivalent and can be measured in the same units. This means that we can talk about a distance of 10^{-3} s (which is ~ 300 km), or a time of 18×10^9 m (equivalent to ~ 1 min). The equation of state (equation (2.20)) then reads $p = w\rho$ and the fluid equation (equation (2.23)) can be rewritten in the form

$$\frac{d}{dt}\left(\rho a^{3(1+w)}\right) = 0, \tag{2.24}$$

with the equivalent solutions

$$\rho = \rho_0\left(\frac{a}{a_0}\right)^{-3(1+w)} \qquad \text{or} \qquad \rho = \rho_0 \mathrm{e}^{-3(1+w)N} \tag{2.25}$$

where

$$N = \ln\left(\frac{a(t_{\text{after}})}{a(t_{\text{before}})}\right) \tag{2.26}$$

is the number of e-foldings between t_{before} and t_{after}. A higher number of e-foldings means that the size of the Universe has expanded by a bigger

[11] These equations can also be derived from Newtonian mechanics with conservation of energy, and the first law of thermodynamics ($dU = TdS + dW$) (see section 4 [15]).

fraction (e^N), which will become relevant in section 2.6. Equation (2.25) together with the equation of state parameter w for a given fluid gives the evolution of that fluid with the expansion of the Universe through the scale factor.

2.4.1 Simplified solutions in an FLRW Universe

To understand the different eras of evolution in the history of the Universe and the differences between these eras some some simple examples are useful. By simplifying the equations to study each component in turn, their evolution can be studied independently of the other constituents.

Radiation or matter only Universe

Consider the simple case of a flat, negligible-Λ Universe that contains only a single component from which it follows that the Friedmann equation simplifies to

$$H^2 = \frac{8\pi G}{3}\rho, \tag{2.27}$$

This simplistic model gives evolutions in time for the scale factor, Hubble rate and density in the Universe of

$$a(t) = a_0 \left(\frac{t}{t_0}\right)^{2/[3(1+w)]}, \tag{2.28}$$

$$H(t) = \left[\frac{2}{3(1+w)} t^{-1}\right] \quad \text{and} \tag{2.29}$$

$$\rho(t) = \rho_0 \left(\frac{a}{a_0}\right)^{-3(1+w)} = \rho_0 \left(\frac{t}{t_0}\right)^{-2}. \tag{2.30}$$

Note that the evolution of density with time is independent of the equation of state parameter and therefore independent of fluid type. As mentioned in section 2.1 the dominant component in the Universe changes with time, and therefore the evolution of the scale factor and the Hubble rate are dependent on which component is dominant[12].

For a flat Universe with only radiation ($w = 1/3$) or only matter ($w = 0$) the simplified version of the Friedmann equation gives the following evolutions for scale factor and density:

$$a(t) = a_0 \left(\frac{t}{t_0}\right)^{1/2}, \ \rho(t) = \rho_0 \left(\frac{a_0}{a(t)}\right)^4 = \rho_0 \left(\frac{t_0}{t}\right)^2 \quad \text{and} \tag{2.31}$$

$$a(t) = a_0 \left(\frac{t}{t_0}\right)^{2/3}, \ \rho(t) = \rho_0 \left(\frac{a_0}{a(t)}\right)^3 = \rho_0 \left(\frac{t_0}{t}\right)^2, \tag{2.32}$$

[12]For a more complicated model, the evolution is determined by the combination of components.

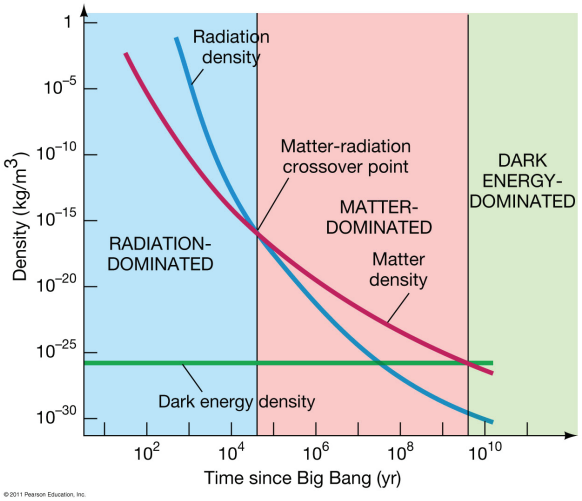


Figure 2.5: Schematic of evolution of density and dominant eras for the three major constituents in the Universe; radiation, matter and dark energy, which are here represented by, respectively, blue, red and green curves and areas. Credit: Pearson Education.

where ρ is given by ρ_r, the radiation energy density, and ρ_m, the matter energy density, respectively. From this it follows that the Universe expands more slowly when filled with radiation only than when filled with matter only, whereas the density evolution is the same in time, but different with scale factor. For a matter filled Universe the energy density in equation (2.31) falls with the increase of volume, here represented in terms of the scale factor by $(a_0/a(t))^3$. The additional factor of $a_0/a(t)$ for the density evolution of a radiation filled Universe in equation (2.32) comes from the fact that the wavelength of the relativistic particles that make up the radiation fluid is redshifted due to the expansion of the Universe, which reduces their energy, and hence the density of the radiation component, by this additional factor.

Following the mentality of a Big Bang model the Early Universe was very hot with the constituents in thermal equilibrium. The relativistic radiation fluid would have dominated the energy content at early time, giving an evolution following equations (2.31) as shown in blue in figure 2.5 for the energy density. As the Universe aged and expanded the temperature of it's constituents dropped and components that were initially relativistic would eventually become non-relativistic (matter) when their thermal energies fell below their rest mass energies. Furthermore, another effect of the expansion, as mentioned already, is that the energy densities of radiation falls faster with

respect to the scale factor than that for matter, thus the contribution from the non-relativistic matter would have become increasingly important. At some point the two fluids contributed an equal amount to the overall energy density of the Universe, when $\rho_r = \rho_m$, a point in time that is given the appropriate name of matter-radiation equality. In the case of the full Big Bang model this occurred at a redshift of $z_{eq} \approx 3400$ [19] as measured by the *Planck* satellite.

After matter-radiation equality the trend of density decrease with the expansion of the Universe would have continued, but due to different rates the matter component would have taken over as the dominant contributor as represented by the red curve in figure 2.5.

The shift between equations 2.31 and 2.32 would not have been instantaneous and there would have been a transition period where both densities were significantly high and had to be combined. The resulting Friedmann equation would have looked like

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}(\rho_r + \rho_m), \tag{2.33}$$

where the relative density of the two components was important, and this ensured a smooth transition from radiation dominated to matter dominated.

Curvature

The cases above describe a flat Universe with no curvature, $k = 0$, and such a Universe will expand asymptotically, with a decelerating expansion that never reaches zero. For the other two curvature states, $k = -1$ and $k = +1$, the Friedmann equation reads

$$\left(\frac{\dot{a}}{a}\right)^2 = -\frac{k}{a^2}, \tag{2.34}$$

as the curvature term becomes dominant, which will eventually happen (in the case of a non-zero curvature) because this term falls off much slower with respect to the scale factor than that for both radiation and matter. For a negative curvature the Universe will, like the zero-curvature case, expand forever, however this expansion will follow $a \propto t$ in a runaway expansion. The positive curvature case is somewhat more spectacular as it causes a deceleration of the expansion rate when it becomes the dominant contribution. This means that the expansion in a Universe with positive curvature ($k > 0$) will eventually grind to a halt and the Universe will recollapse under gravity, ending in another singularity[13] dubbed the Big Crunch.

[13]The Big Bang and Big Crunch are gravitational singularities that arise when the energy density becomes infinite.

2.5 Pillars of the Big Bang theory

There are several pieces of observational evidence that fit with the predictions of the Big Bang model. However, three of them stand out and are often referred to as the pillars of the Big Bang model:

Expansion of the Universe The recession velocity of a galaxy is computed from its observed redshift based on the galaxy's observed spectrum and line emissions. Using Hubble's law the distance to galaxies can be computed, which indicate that the Universe is expanding[14].

Primordial nucleosynthesis The predicted abundances of the lightest elements from the Big Bang theory agree well with the observed abundances in the Universe [8, 9, 29, 30, 31].

Existence of the CMB The properties of the cosmic microwave background radiation are very close to that predicted from theory [11, 12, 32, 33].

2.6 Extensions to the Big Bang theory

As mentioned in section 2.3, observations indicate that the curvature of the Universe is very close to flat with the most recent measurements by the *WMAP* and *Planck* satellites giving a total density fraction of $\Omega_{\text{tot}} = 1.0027^{+0.0038}_{-0.0039}$ [18] and $\Omega_{\text{tot}} = 1.0005^{+0.0065}_{-0.0066}$ [19], respectively. However, a nearly flat non-zero curvature Universe creates some inconsistencies in the Big Bang model. When taking into account only the energy densities as a result of radiation and visible (baryonic) matter the energy density ratios, Ω_r and Ω_b, do not even get close to adding up to unity and are found to account for only about 5% of the critical density[15].

To ensure flatness additional components are introduced to the Big Bang model, the first of these, dubbed dark matter, makes up slightly more than a quarter of the energy content required and therefore dominates matter evolution. Indirect detections of dark matter have been made through analyses of the virial mass in clusters [34, 35], the rotation curves of spiral galaxies [36] and through mass estimates from gravitational lensing [37]. Furthermore, dark matter is heavier than ordinary matter and therefore decouples and starts clumping earlier, thereby setting up potential wells for ordinary matter to fall into. Without the existence of dark matter the ordinary matter would not have had sufficient time to produce the large scale structures

[14] Supernova observations by Perlmutter et al. [26], Schmidt et al. [27] and Riess et al. [28] indicate that this expansion is accelerating, for which they received the 2011 Nobel Prize in Physics.

[15] Radiation contributes only a very small fraction to these 5%, of the order of $\Omega_r \sim 10^{-6}$.

observed. However, the nature of the particle(s) associated with dark matter still eludes scientists.

The second component added to bridge the gap towards a flat Universe goes by the name dark energy and should account for nearly 70% of the energy content of the Universe today. It is also the leading contender for causing the Universe's accelerated rate of expansion. This acceleration comes about as a consequence of the dark energy having an equation of state parameter smaller than zero meaning that it has a negative pressure, causing space to expand at an ever increasing rate. Even though dark energy is the dominant constituent of the Universe today (provided it exists), we know even less about its nature than we do about that of dark matter, which basically means that we have no clue what makes up roughly 95% of the Universe! These two components give us the ΛCDM model, which build on the Big Bang model, but also contains dark energy, represented by Λ, and cold dark matter, where cold refers to the dark matter being non-relativistic.

For a flat Universe containing only dark energy the Friedmann equation simplifies to

$$H^2 = \frac{\Lambda}{3}, \tag{2.35}$$

which is called the de-Sitter solution, and gives an exponential evolution for the scale factor as

$$a(t) = a_0 \mathrm{e}^{\sqrt{\Lambda/3}\, t}. \tag{2.36}$$

The total energy content of the Universe is today dominated by this mysterious dark energy, but this was not always the case as the Universe at early hot times was radiation dominated, and somewhat later as the Universe cooled it was dominated by matter. The energy density of the dark energy is constant in time, given by

$$\rho = -p = \frac{\Lambda}{8\pi G}, \tag{2.37}$$

and will eventually become dominant since the energy densities of the other constituents decrease with time, dropping as a^{-4} for radiation and a^{-3} for matter. Figure 2.5 shows the evolution of the Universe in terms of these three main constituents where green corresponds to dark energy. As can be seen it is only recently that the densities of radiation and matter fell below the energy density of dark energy, which therefore came to dominate the evolution of the Universe. The fact that the dark matter – dark energy phase transition happened so recently (at $\sim 10^9$ yrs compared to the present age of $\sim 10^{10}$ yrs) is a slight worry and is often referred to as the coincidence problem.

As mentioned in section 2.5, the theoretical framework of the Big Bang theory provides explanations for a variety of astrophysical observations such as the three pillars; the expansion of the Universe, the abundances of lighter elements through Big Bang nucleosynthesis and the existence and properties of the CMB radiation. Further evidence exist in the shape of the existence, sizes and numbers of astrophysical objects; structure on large scales and much more. However, some observations cannot be explained without having to introduce some new theoretical ideas, being mainly:

Horizon problem. According to general relativity no information can travel faster than the speed of light, thus regions of the sky that are further apart than this limit, which is referred to as the particle horizon, cannot have shared information and will therefore not be in equilibrium. In the case of the CMB the angular size of this horizon is of the order of 1° and therefore regions in the sky that are further apart than this should not be causally connected. However, the CMB signal that we observe shows that even regions in opposite directions are in thermal equilibrium.

Flatness problem. The classification of the spatial curvature of the Universe is a consequence of whether its energy density is smaller than, equal to or larger than the critical energy density of the Universe, and is referred to as being either an open, flat or closed Universe, respectively. If there is any deviation from the critical density it will grow with time following $[\Omega_{\text{tot}} - 1] \propto a^2 \propto t$ in the matter dominated era or $[\Omega_{\text{tot}} - 1] \propto a \propto t^{2/3}$ in the radiation dominated era, thereby taking the Universe further and further from flatness. However, cosmological observations indicate that the spatial curvature of the Universe today is very close to flat, which means that it must have been exceptionally close to flat in the past indicating an uncomfortable level of fine-tuning.

Monopole problem. Grand Unified Theories (GUTs) unify the strong, weak and electromagnetic regimes and postulate the existence of magnetic monopoles in the Universe, which should have been produced at observable densities during the very early hot stages. However, none of the searches for these monopoles have ever been successful so either the GUTs are completely flawed or a piece of the puzzle is missing.

A popular way to solve these three problems, which was put forward by Guth [38], Sato [39], Linde [40], and Albrecht and Steinhardt [41] in the early 80s, is to introduce an epoch called cosmological inflation to the existing ΛCDM theory, at a time when the Universe was around 10^{-35} seconds old. During this epoch the Universe developed a negative pressure causing it to go through an extremely rapid[16] exponential expansion in which it went through at least 50 e-foldings, corresponding to a factor of e^{50} for the scale factor [42]. This expansion would have caused the regions of space that were in causal contact with one another to become extremely large, even to the extent that the entire observable Universe today was in causal contact before inflation. With such an extreme epoch of expansion the CMB radiation would have originated from a region of space smaller than that required to be in causal contact before the inflationary period, thereby solving the horizon problem. Furthermore, the inflationary expansion would have stretched space causing a smoothing out of any deviations from flatness, to the levels required to match that observed today. Similar to the reasoning to allow for the entire observable Universe today to have been in causal contact before inflation, the monopole problem is explained by considering that the expansion could have lead to the removal of all monopoles from our observable Universe.

In addition to solving these three problems, it is also theorized that during the inflationary era quantum fluctuations occurred, which evolved to perturbations in the otherwise smooth distribution of matter [43, 44, 45]. These tiny fluctuations are postulated to be the seeds for large scale structure since overdensities and underdensties grow larger through gravitational attraction and the lack there of. Eventually the overdensities would have grown so large that gas clouds would have collapsed and the high temperatures in clouds with extremely dense cores would have caused them to ignite, thereby producing the first stars 12 billion years ago. However, based on fundamental physics there is no mechanism to cause the exponential expansion during inflation or an explanation for how it is ended. On the other hand an impressive number of inflationary theories exist involving scalar, vector and tensor fields, slow roll, chameleon fields, multiple fields and much more.

[16] The expansion was faster than the speed of light.

Chapter 3

CMB observations and analysis

3.1 CMB radiation

The surface of the cosmic microwave background radiation is presently the furthest back in time it is possible to probe with an astrophysical source, and studying the CMB signal is therefore often referred to as looking at a baby picture of the Universe. The CMB photons originate from a time roughly 380 000 years after the Big Bang, as a result of a phase transition in the Universe, known as the recombination or decoupling era. Before this time the photons were essentially trapped as they continually scattered off free electrons and ions in the baryonic plasma with a temperature of over 3000 K. Therefore, the photons and baryons were in thermal equilibrium, and the radiation followed the spectrum of a black body[1]. As the temperature decreased, the first atoms formed and later the photon scattering ceased as the ionization fraction[2] became negligible because most electrons were trapped in atoms. Due to the lack of free electrons to interact with the photons were free to stream through the Universe with a mean free path exceeding the known Universe, and the radiation retained its black body spectrum. These photons are the same as those reaching us today and can be thought of as coming from a spherical shell centered on the observer, at a distance of the time since recombination. This is illustrated in figure 3.1 which shows the Milky Way galaxy in the centre of a set of spheres (cut in half), where the radius of a sphere is proportional to the time of emission (look-back time), and the outermost sphere represents the surface of last scattering from which the CMB photons originate.

[1] A black body is an object that is both a perfect absorber and re-emitter of electromagnetic radiation.

[2] The ionization fraction (X) is the number of free electrons (n_e) to baryons (n_b) in the Universe, $X = n_e/n_b$.

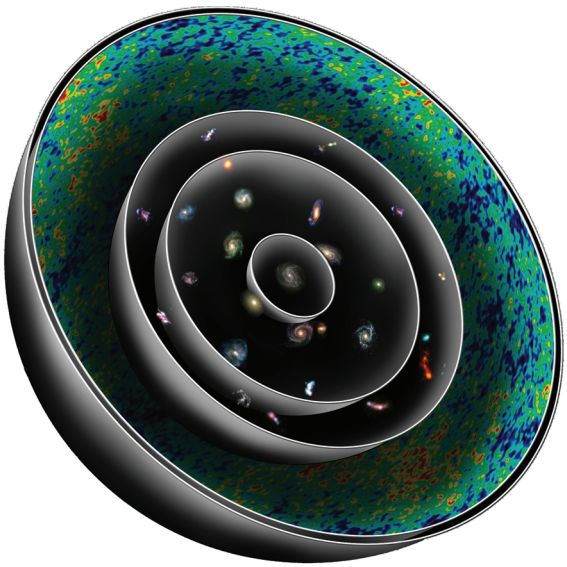


Figure 3.1: Schematic of a set of cosmic spheres as view from the Milky Way galaxy, in the center of the spheres. The radius of a sphere represents the look back time, which is the time since the light from its sources was emitted and therefore gives a measure of distance. The CMB surface is the outermost sphere as this is the furthest back in time it is possible to look with an astrophysical source, and all other spheres contain foreground sources. Credit: Nicolle Rager Fuller (`http://viewfromthecenter.com`).

As they traveled through the Universe, CMB photons were scattered and distrupted by structures along the way. Examples of such disruptions are the integrated Sachs-Wolfe (IWS) effect [46] and the Sunyaev-Zel'dovich (SZ) effect [47, 48] that distort the CMB radiation due to gravitational redshifting and through inverse Compton scattering in galaxies, respectively. For further details and the most recent results, courtesy of the *Planck* Collaboration, see the 2013 release papers [49] and [50]. However, these effects only result in small perturbations, and the most drastic change to the radiation is the linear temperature shift of the CMB photons as they travel through the Universe. As mentioned in section 2.4.1, the energy density evolution of relativistic particles exhibit an extra redshifting factor due to the expansion of the Universe as the space through which the photons travel is stretched. The amount of stretching is directly related to the redshift, as seen in equation (2.3), and using the redshift-scale factor relation from equation (2.4)

the photon wavelength at any time, $\lambda(t)$, can be computed through

$$\lambda(t) = \lambda(t_{\rm em})\frac{a(t)}{a(t_{\rm em})}, \tag{3.1}$$

where $t_{\rm em}$ is the time of emission, which in the case of CMB photons is the time of recombination. From this it follows that the wavelength after the Universe has doubled its size is twice as long as when emitted, and furthermore that the CMB radiation reaching us today have wavelengths that are roughly 1100 times larger than at recombination as this is the amount by which the Universe has been scaled up since that time.

3.1.1 Black body curve

When discussing the CMB radiation it is more common to talk about its temperature T, as opposed to its wavelength λ, and results from CMB experiments are usually quoted in terms of temperature. By assuming that the CMB radiation has a black body spectrum, the intensity of the CMB photons, $I(\nu, T)$, follow a distinct distribution according to their temperature T and frequency ν, which is given by Planck's law,

$$I(\nu, T) = \frac{2h\nu^3}{c^2}\frac{1}{{\rm e}^{h\nu/kT} - 1}, \tag{3.2}$$

where h, c and k are the Planck constant, speed of light in vacuum and Boltzmann constant, respectively. This is justified since in the early 90s the COsmic Background Explorer (*COBE*) [51] measured the CMB intensity and found it to fit a black body curve to stunning precision with a temperature of $2.725 \pm 0.002\,{\rm K}$ [52], as illustrated in figure 3.2. The errors on the intensity measurements in this plot are smaller than the thickness of the curve, leaving little doubt as to the blackbody nature of the CMB radiation. More recent observations have improved the measurement of the temperature to $2.72548 \pm 0.00057\,{\rm K}$ [53] by combining the results from *COBE*'s FIRAS detector and the 5-year *WMAP* data.

For radiation with high frequencies a simple relationship exists to relate the peak wavelength, $\lambda_{\rm peak}$, in Planck's law to the temperature of the black body radiation, T, through

$$T\lambda_{\rm peak} = 2.898 \cdot 10^{-3}\,{\rm mK}, \tag{3.3}$$

which is known as Wien's displacement law. For a black body with a given temperature, measurements at the peak wavelength always yield a higher intensity than measurements at longer or shorter wavelengths, but the temperature measurement is identical. Furthermore, CMB measurements are usually performed at more than one wavelength to provide information of the shape of the intensity spectrum and to ease foreground cleaning.

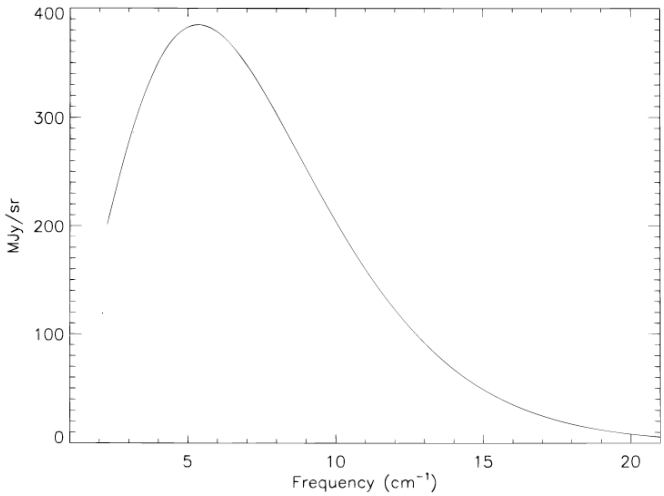


Figure 3.2: Black body spectrum from the full data set of *COBE* where the errors on the measurements are smaller than the line thickness. Source: Fixsen et al. [54].

3.2 CMB observations

The analyses in this thesis are mainly performed on observational data of the cosmic microwave background radiation, which is observed in a region of the electromagnetic spectrum that is invisible to the human eye. The observational regime for the CMB lies in the radio range[3] with frequencies between 10 GHz and 1000 GHz. However, the CMB radiation's presence can also be viewed directly on an old analogue television that is tuned away from a broadcasting channel, thereby leaving the "white snow". About 1% of this snow is due to interference by the photons of the CMB!

As mentioned, the first observational evidence for the existence of the cosmic microwave background radiation was made by Penzias and Wilson [12] in 1964 and since that time a great number of experiments have been designed to study its properties either from the ground, on balloon-borne crafts or from space. Examples of ground-based experiments are the Cosmic Background Imager (CBI) [55, 56], Degree Angular Scale Interferometer (DASI) [57], Arcminute Cosmology Bolometer Array Receiver (ACBAR) [58], Atacama Cosmology Telescope (ACT) [59] and the South Pole Telescope (SPT) [60]. Some of the balloon-borne experiments that have been carried out include the Balloon Observations Of Millimetric Extragalactic Radiation ANd

[3] In this range the CMB signal is reasonably high compared to foregrounds.

Geophysics (BOOMERanG) [61], Millimeter Anisotropy eXperiment IMaging Array (MAXIMA) [62, 63] and Archeops [64]. The great advantage of ground-based and balloon-borne experiments is that compared to space-based experiments they are reasonably cheap and easy to repair, although repair is often not possible once balloon experiments have been launched. These types of experiments are primarily used to study features of the CMB on small scales, from dry and high altitude sites, where ground-based experiments are limited to the surface of the Earth. These experiments have to take into account the changing conditions at their sites/flight plans, as well as having to deal with noise caused by the atmosphere, which is less of a problem for balloon experiments since they fly above Earth's surface. Ground and balloon experiments cannot be used to study large scale phenomenon since Earth gets in the way and the experiments usually limit their areas of investigation to small patches on the sky, which are studied thoroughly. For a full-sky survey the experiment has to be launched into space, which is considerably more expensive.

The first space-based satellite to observe the full sky CMB radiation was the COsmic Background Explorer (*COBE*) [51] which orbited Earth in a sun-synchronous orbit[4] in the early 90s and is by many regarded as the start of high precision experimental cosmology. The first release from the Far Infrared Absolute Spectrophotometer (FIRAS) instrument on *COBE* unveiled a CMB radiation with a perfect black body spectrum [65] across the whole sky, but the Differential Microwave Radiometer (DMR) instrument revealed that the CMB temperature is not completely isotropic but also exhibits some temperature anisotropies of the order of $\Delta T/T \sim 10^{-5}$ [66]. These temperature anisotropies were later studied by the ground and balloon experiments already mentioned, and by two other space-based experiments.

3.2.1 The *WMAP* satellite

In 1995 the Microwave Anisotropy Probe (*MAP*)[5] was proposed to the National Aeronautics and Space Administration (NASA) as a space-based satellite mission to study the temperature anisotropies of the CMB radiation, that were found by the *COBE* team at a level of $\Delta T/T \sim 10^{-5}$ [66, 67, 32]. The project was approved for development in 1997, launched in June 2001 and headed for a special destination to conduct its observational run. For the first time a satellite would travel out to the L2 Lagrange point to perform its measurements, which is one of five such Lagrange points that allow a small body to orbit together with two larger ones, as can be seen in figure 3.3. L2 is located along the axis joining the Sun and Earth, at an orbit exterior to

[4] A sun-synchronous orbit is a geocentric orbit around Earth in the plane that is perpendicular to the direction to the Sun, that will change with respect to Earth's surface as Earth orbits the Sun.

[5] Website: `http://map.gsfc.nasa.gov/`

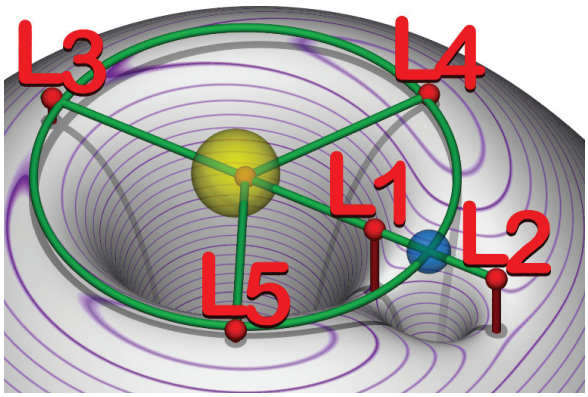


Figure 3.3: Illustration showing the positions of the five Lagrange points (red spheres) for the Sun-Earth system in the combined gravitational potential of the Sun (yellow sphere) and Earth (blue sphere), where the purple lines represent field lines. Lagrange points L1, L2 and L3 are unstable as they reside at saddle points, whereas L4 and L5 are stable points at the peaks of the field. Source: Wikipedia.

Earth, where there is a saddle point in the combined gravitational fields of the Sun and Earth. The L2 point, together with L1 and L3, are classed as unstable Lagrange points since any significant perturbation from the saddle would cause the object residing there to drift from the original position, and the satellite required some tweaking to keep it in the saddle. However, L2 is nevertheless an excellent location for astrophysical observations as it allows for easy shielding from the Sun, Moon and Earth while still being in close proximity to Earth[6] allowing for relatively fast communication. Lagrange points L4 and L5 are classed as stable Lagrange points due to the fact that they occur at peaks in the combined gravitational field, and objects placed at these points will return without requiring on board adjustments after small perturbations. However, their distance from Earth is not ideal for communication purposes. Furthermore, a number of small objects, called Trojans, reside at L4 and L5 making it a somewhat hostile environment.

After a few months of travel the *MAP* satellite arrived at L2 in August of 2001 and started mapping the temperature of the microwave sky. The data from the first year of observations was released in February 2003 and around the same time the satellite was renamed the Wilkinson Microwave Anisotropy Probe (*WMAP*) in memory of David Todd Wilkinson[7], who had

[6] L2 is approximately $1.5 \cdot 10^9$ m from Earth, which is 1 hundredth of the distance to the Sun or about 4 times the Earth-Moon distance.

[7] Dr. Wilkinson lost the fight with cancer in the Autumn of 2002.

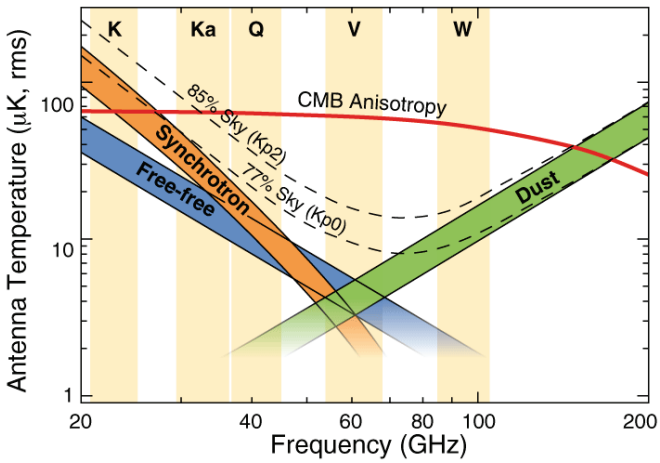


Figure 3.4: The five frequency bands for observations with the *WMAP* satellite overlayed on the frequency dependence of the CMB anisotropy and the three major foregrounds; synchrotron, free-free and dust. Source: NASA / *WMAP* Science Team.

played an important role in the mission and in the field of CMB analysis in general. The satellite was originally scheduled to collect data for only 2 years but was repeatedly granted extensions resulting in a total of 9 years of data collection. The final day of data collection was the 19th of August 2010 and *WMAP* was finally decommissioned[8] in late October of the same year. The major scientific achievements of the *WMAP* satellite was to confirm *COBE* findings, indicating a nearly perfect black body spectrum, and the investigation of the temperature anisotropies to a much better resolution of below 1° scales. The latter of these enabled the construction of the CMB power spectrum quantifying the amount of clumping on different scales.

In order to understand the scientific achievements of the *WMAP* satellite let us look at some of the specifics regarding the data collection and processing. The temperature anisotropies of the microwave sky was observed by using differential microwave radiometers to measure the difference in temperature between two points on the sky. The CMB signal originates from the earliest visible time in the history of the Universe and there are plenty of contaminants between this background signal and us, such as point sources and various foreground emissions from the Milky Way Galaxy. These contaminants have to be removed in order for the CMB sky to be revealed, which

[8]Decommissioned means that *WMAP* ceased to communicate, and was set to orbit the Sun out beyond L2.

requires observations at multiple frequencies and the use of component separation techniques to separate the CMB signal that we are interested in from the various foregrounds. Therefore, the *WMAP* satellite was designed to observe the microwave sky at 5 different frequency channels, going by the names K, Ka, Q, V and W, covering a frequency range between 23 and 94GHz, as seen in figure 3.4. The *WMAP* requirements were to measure the temperature anisotropies with an angular resolution of at least 0.3° and a sensitivity of 20 μK per 0.3° square pixel. The sensitivity and resolution requirements were surpassed and allowed for an angular power spectrum which could reach to as small scales as a multipole moment of $\ell \sim 1200$. The scanning strategy adopted by *WMAP* was to rapidly scan the sky with a slow precession, meaning that the same spot was covered at different angles and with different intervals between scans. Having more than a single scan of the full sky, which each took six months to complete, allows for better stability checks. *WMAP* produced a total of 18 full-sky maps before being decommissioned, and the data set after every odd number of years of observation were made publicly available on LAMBDA[9].

3.2.2 The *Planck* satellite

What is now known as the *Planck* mission started out as two independent proposals to study the CMB temperature anisotropies, that were both made to the European Space Agency (ESA) in 1993 for entry into the Horizon 2000 Scientific Program. Due to the similarities of the two missions they were joined up as COBRAS/SAMBA[10], but after being approved in 1996 it was renamed *Planck* in honor of the German physicist Max Planck who contributed to a variety of theoretical developments in physics and was one of the founding fathers of quantum theory.

The early stage of the *Planck* mission was being developed at the same time as the *WMAP* satellite, which was also designed to study the CMB temperature anisotropies but which would launch much earlier. What was the scientific merit of developing and launching *Planck* after the success of the *WMAP* mission? Even though the two satellites were designed to investigate the same signal, the satellite design, scanning strategy and, most importantly, the sensitivity and resolution were very different. As mentioned in section 3.2.1 *WMAP* could accurately measure the CMB temperature to produce an angular power spectrum up to a multipole of $\ell \sim 1200$ whereas *Planck* was designed to be able to measure the CMB temperature to higher sensitivity, and therefore be able to go to much smaller scales than *WMAP*,

[9] Legacy Archive for Microwave Background Analysis,
website: `http://lambda.gsfc.nasa.gov/`

[10] From combining the abbreviations of the two original proposals; COsmic Background Radiation Anisotropy Satellite and SAtellite for Measurement of Background Anisotropies, respectively.

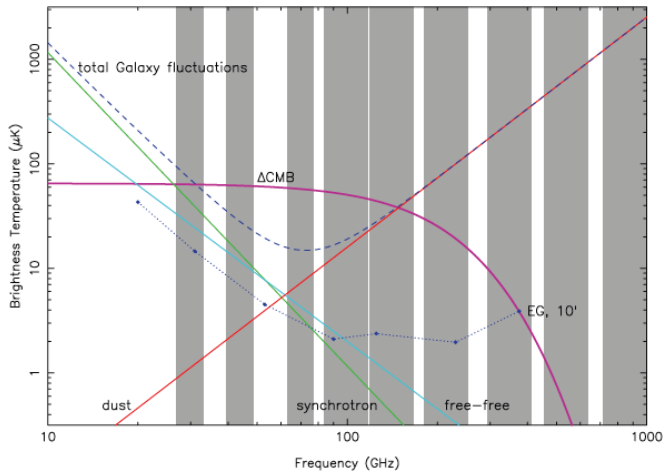


Figure 3.5: Brightness temperature of sources with the nine frequency bands of the *Planck* satellite overlayed, where the three lowest frequency channels are contained in the LFI instrument, and the rest in the HFI instrument. The CMB anisotropy signal and the signal from the three major foregrounds; synchrotron, free-free and dust, are shown, together with the total galaxy fluctuation and the expected fluctuation level of foreground radio sources on a $10'$ scale. Source: *Planck* blue book.

reaching at least $\ell \sim 2000$. This increased sensitivity had to be accompanied by an increase in detector sensitivity and a larger frequency range, which took time to develop and test. The development period for *Planck* was therefore significantly longer than for *WMAP* and it was not launched until May 2009, eight years after the launch of *WMAP*. *Planck* had a total of 9 frequency channels as can be seen in figure 3.5, that are split between two instruments with completely different detector technologies. The Low Frequency Instrument (LFI) was made up of a set of radio receivers detecting at the three lowest frequency channels, centered at 30, 44 and 70 GHz. The High Frequency Instrument (HFI) covered the remaining channels, centered at 100, 143, 217, 353, 545 and 857 GHz, and consisted of bolometric detectors. In order to reach the temperature sensitivity of the order of a millionth of a degree the *Planck* detectors required to be cooled to 0.1 K by the accompanying cooling system, which further complicated the satellite design.

The L2 Lagrange point was also chosen to be *Planck*'s home for the duration of the mission, where it arrived in July 2009 and promptly started taking measurements. In January 2012 after 2.5 years of observations five full-sky

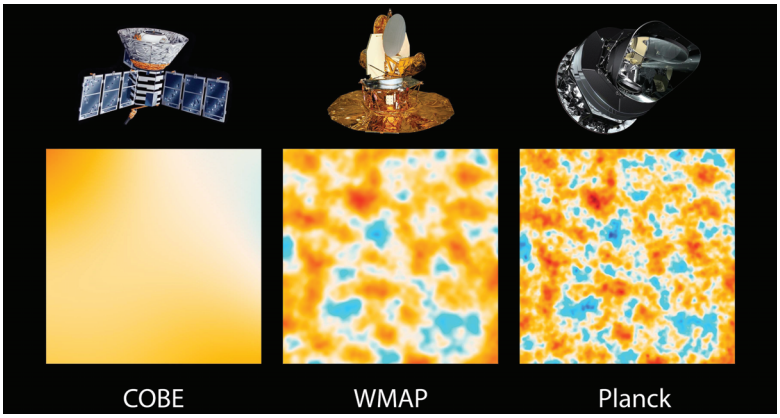


Figure 3.6: Resolution comparison and images of, from left to right, *COBE*, *WMAP* and *Planck*, where the patches show the same 10°-by-10° area of the satellites' all-sky maps. Credit: NASA/JPL-Caltech/ESA.

surveys had been completed and the cooling system ran out of coolant, as expected, and HFI ceased to collect data. However, up to this point the mission had been a great success, compared to the minimum requirement for a successful mission of two full-sky surveys. The LFI instrument can operate at slightly higher temperatures than HFI and continued to collect data until early August 2013. After completing the final calibration measurements *Planck* was maneuvered away from L2 and finally decommissioned on the 23rd of October 2013. On the 21st of March 2013 the *Planck* Collaboration released the first data set and official *Planck* results[11] from analyses of the temperature measurements for the nominal mission comprising of 15.5 months of observations. The next release in 2014 will contain the full data set, consisiting of 30 months of HFI observations and 48 months of observations with LFI, and in addition the first polarization data and results.

Figure 3.6 depicts the resolutions of the satellite observations and the improvement with time, where the panels from left to right show the same 10° by 10° patch of the full CMB sky for the *COBE*, *WMAP* and *Planck* satellites. The large scale structure in the *COBE* patch is also visible in the patches of *WMAP* and *Planck*, and the resolution improvement is astounding. The hot and cold spot structure observed with *WMAP* are clearly visible in the *Planck* patch, but the improved resolution from roughly 15 arcminutes to about 5 arcminutes allows for perturbations on much smaller scale to be observed in the *Planck* data.

[11] The *Planck* Legacy Archive and official papers are available at `http://www.sciops.esa.int/index.php?page=Planck_Legacy_Archive&project=planck`.

3.3 Modeling the CMB

The most basic way to model CMB data, $\boldsymbol{d}$, is given by

$$\boldsymbol{d} = \boldsymbol{s} + \boldsymbol{n}, \tag{3.4}$$

where $\boldsymbol{s}$ is the true CMB signal and $\boldsymbol{n}$ the instrumental noise. Both of these variables are assumed to be Gaussian variables with zero mean and covariance matrices $\boldsymbol{S}$ and $\boldsymbol{N}$, respectively. This model assumes that the only non-CMB signal is the noise contribution from the instruments performing the observations. However, reality is a lot more complicated, where observational constraints and contaminants also have to be take into account, which gives rise to a more complicated model,

$$\boldsymbol{d} = \boldsymbol{PBs} + \boldsymbol{n} + \sum_{i=1}^{n_f} \boldsymbol{f}_i, \tag{3.5}$$

where $\boldsymbol{P}$ and $\boldsymbol{B}$, the pixel window and instrumental beam are observational constraints that effect the CMB signal, and $\boldsymbol{f}_i$ are a set of n_f foreground contaminants that are unavoidable when performing CMB observations.

Instrumental beam. The finite width of the detection beam for a detector horn causes a smoothing effect called the instrumental beam, or for short just the beam. A horn pointed towards a spot in the sky will not just pick up the CMB temperature from this single point but also sample the signal from an infinite number of points in a certain area around the pointing direction. In other words the CMB signal that the detector picks up in a given direction is impure as this signal is in fact a smoothed out signal from the area covered by the detector beam.

Foreground contamination. Foregrounds are the bane of cosmologists; they are sources that obstruct our view of the CMB and have to be removed in order to access the underlying CMB signal. These are therefore good examples of the proverb that 'one man's trash is another man's treasure', since these sources are the research topics of other astrophysicists. The most obvious source of contamination is the Milky Way galaxy that contributes a variety of different foregrounds, which each have a frequency dependency, intensity and volume it originates from. Amongst the galactic foregrounds we find free-free, dust, spinning dust and synchrotron emissions, in addition to line emission from different elements, for example from CO and HI. The maps[12] in figure 3.7 show examples of galactic foregrounds from the 2013

[12] A map of the sky (e.g. of the CMB, foregrounds, the Milky Way galaxy or point sources) is a projection of the two-dimensional field on the spherical surface onto a flat representation, similar to the flattening of the spherical map of Earth onto a flat representation.

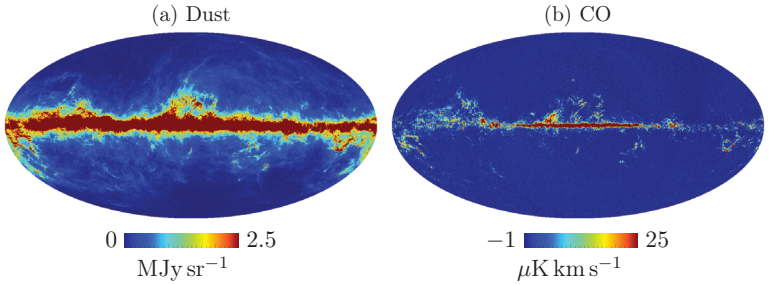


Figure 3.7: Galactic foreground emission from dust and CO with accompanying color bars that indicate emission intensities which are stronger along the galactic plane than at higher latitudes. Dust emission stretches to higher latitudes than that of dust, but they have many emission regions in common as they both trace the arm structure of the Milky Way. Source: *Planck* Legacy Archive.

Planck release in the form of emission from dust (a) and CO (b), with accompanying color bars indicating their range of emission intensity. The dust emission clearly originates in regions that extend further out from the galactic disk than that of CO. On the other hand, since the emissions originate in the galactic arm structure there are a lot of overlapping regions that are especially obvious at higher latitudes. For example, there are similarities in the structures of the Orion-Eridanus Bubble slightly below the galactic plane on the far right of the maps, as well as in features of the Galactic Haze above the central region.

In addition to these galactic emission foregrounds, there are a number of other foregrounds, for instance the extra-galactic cosmic infrared background (CIB) that is emitted by galaxies in the early Universe, as well as a great number of point sources such as stars in the Milky Way galaxy, and other galaxies[13]. The solar system also contributes to the foregrounds and such local contaminants include dust emission from the solar system, called the Zodiacal light, as well as planets and objects in the solar system. In the case of ground-based or balloon-borne experiments Earth and its atmosphere also add to the contamination of the CMB signal, which can be removed by observing from space. However, the rest of the foreground problems cannot be solved by simply going "elsewhere"[14] to perform the measurements and therefore have to be dealt with in the post-processing, which will be discussed further in chapter 5.

[13]Obviously only objects that are old enough to have witnessed the passing of the CMB photons arriving at us today will turn up as foregrounds.

[14]Hypothetically, "elsewhere" would entail traveling out of the Milky Way galaxy and performing the CMB measurements far enough away from it and other galaxies for their foreground contributions to be negligible. This would not deal with the CIB, which is a universal background signal from after the end of the "dark ages".

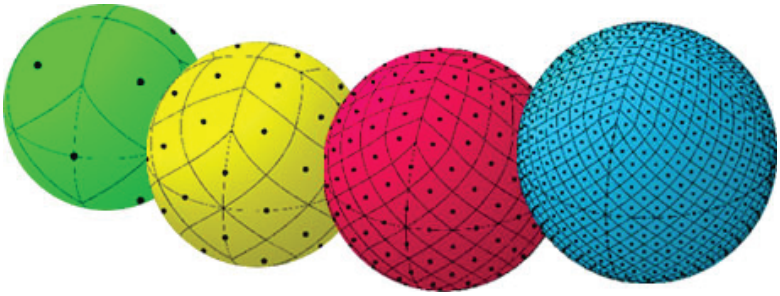


Figure 3.8: Refinement scheme for HEALPix showing pixelization at $N_{\rm side}$s of 1, 2, 4 and 8 for the spheres from left to right.

Pixel window. Lastly we arrive at the pixel window effect, which can most easily be explained after a discussion of the division of the CMB sky into a finite number of pixels. The most commonly used pixelization scheme in the cosmological community is the Hierarchical Equal Area isoLatitude Pixelization (HEALPix) [68] scheme which comes with an accompanying software package[15], and was used for the sky maps of both the *WMAP* and *Planck* data.

This pixelization regime has several levels, labeled by their $N_{\rm level}$ value which starts at 1 and increases by an increment of one for each refinement. The first level has the lowest resolution and contains only 12 pixels over the whole sphere, and each subsequent level contains more pixels, thereby giving an increased resolution with each level, as can be seen in figure 3.8 for $N_{\rm level}$ of 1, 2, 3 and 4 represented by the green, yellow, red and blue spheres, respectively. Two other quantities are also commonly referred to when describing different pixelization levels, namely the $N_{\rm side}$ and $N_{\rm pix}$ values for a level. These represent the number of sub-pixels along one side of any of the 12 minimum resolution pixels and the total number of pixels over the whole sky. For each level of improved resolution every pixel of the previous level is subdivided into two along each edge giving four higher level pixels. It follows that the higher level $N_{\rm side}$ and $N_{\rm pix}$ are found by multiplying that of the previous level by two and four, respectively. These relationships are summarized through

$$N_{\rm level} = \{1, 2, 3, 4, 5....\} \tag{3.6}$$
$$N_{\rm side} = 2^{N_{\rm level}-1} \tag{3.7}$$
$$N_{\rm pix} = 12N_{\rm side}^2 = 12(2^{N_{\rm level}-1})^2. \tag{3.8}$$

[15] HEALPix is downloadable from http://healpix.sourceforge.net/.

'Equal Area' implies that all `HEALPix` pixels at a certain resolution level have the same area, given by

$$A_{\text{pix}}(N_{\text{side}}) = \frac{4\pi}{N_{\text{pix}}} \text{ steradian.} \tag{3.9}$$

Furthermore, 'isoLatitude' entails that the centers of all the pixels are placed along a limited number of rings, $N_{\text{iso-rings}}$, of constant latitude, with an increasing number of such rings with increased resolution. This division is given by

$$N_{\text{iso-rings}} = 4N_{\text{side}} - 1, \tag{3.10}$$

and can be seen in figure 3.8 for the four lowest level pixelizations, which have 3, 7, 15 and 31 such isolatitude rings of pixel centers, respectively.

The representation of the CMB field on the sphere as a pixelized map of this kind causes a certain amount of smoothing as the value at each pixel represents the average signal the scans pick up from the area of this pixel, similar to the smoothing caused by the beam. The pixel window is used to keep track of the necessary correction per multipole for the different pixel refinements.

3.4 The CMB power spectrum

After disentangling the CMB signal from the contaminants the resulting map can be analyzed to ascertain the properties of the Universe we live in, such as estimating the best-fit shape of the angular power spectrum, and evaluating which model and parameter that best describe the Universe. The CMB map published by the *Planck* Collaboration in the nominal release is shown in figure 3.9 where the area inside a 3% mask has been removed and subsequently filled in by a constrained Gaussian realization. As mentioned in section 2.4 the Universe is assumed to follow the Copernican principle by being isotropic and homogeneous, and further, that the CMB signal is also assumed to be isotropic and homogeneous. The CMB is essentially a scalar field on a sphere centered on us, and such a field can be split into into components on different scales by using spherical harmonics[16], $Y_{\ell m}$, through

$$\boldsymbol{s} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}, \tag{3.11}$$

where $a_{\ell m}$ are the spherical harmonic coefficients and each multipole, ℓ, has $2\ell + 1$ solutions. The $\ell = 0$ and $\ell = 1$ solutions represents the CMB monopole and dipole, which are the average CMB temperature T and the

[16]The spherical harmonic representation is the spherical cousin of the Fourier series.

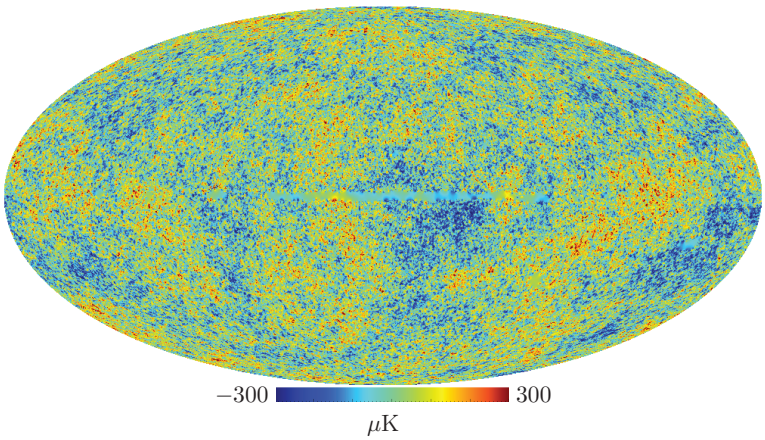


Figure 3.9: The *Planck* Collaboration's CMB temperature map from the nominal 15.5 months of observations, which has been smoothly filled in inside a 3% mask.

dipolar structure caused by our motion with respect to the rest frame of the CMB radiation. Multipoles with $\ell > 1$ correspond to deviations from the average CMB temperature and larger ℓ represent anisotropies at smaller scales[17], where for instance the quadrupole ($\ell = 2$) and octopole ($\ell = 3$) solutions corresponding to scales with two and three waves on the sphere. This means that a temperature anisotropy with angular size α corresponds to perturbation multipole

$$\ell \approx \frac{180^\circ}{\alpha}. \tag{3.12}$$

In order to analyze the CMB temperature fluctuation, $\Theta(\hat{\boldsymbol{n}})$, the monopole and dipole parts of the signal are removed and can be treated as foreground components. This leaves the temperature anisotropies from $\ell \geq 2$, which can be written as

$$\Theta(\hat{\boldsymbol{n}}) = \frac{\Delta T(\hat{\boldsymbol{n}})}{\langle T \rangle} = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{\boldsymbol{n}}) \qquad \text{with} \tag{3.13}$$

$$a_{\ell m} = \int Y^*_{\ell m}(\hat{\boldsymbol{n}}) \Theta(\hat{\boldsymbol{n}}) d\Omega, \tag{3.14}$$

where $a_{\ell m}$ is integrated over all space. The fluctuations are assumed to be Gaussian, which means that the $2\ell + 1$ $a_{\ell m}$ coefficients are statistically both

[17] The solution with multipole number ℓ is that with ℓ complete waves on the sphere, causing a superposition with 2^ℓ poles, giving it the name 2^ℓpole (in latin).

isotropic and independent with a mean of zero and variance given by

$$C_\ell = \langle |a_{\ell m}|^2 \rangle, \tag{3.15}$$

called the power spectrum. C_ℓ gives the amplitude of the temperature perturbations on different scales in the Universe. In other words it shows the amount of clumping on different scales over the whole sky. The shape of the power spectrum is related to how the different modes evolved between inflation and recombination, and depends on when they entered the horizon and their state when they were frozen out at recombination. Before recombination, matter and radiation were tightly coupled, which included their evolutions, however at recombination this coupling was broken and they evolved independent of each other. In the case of matter this meant that the perturbations were free to grow through gravitational collapse to produce large scale structure, while for radiation the perturbations were almost constant as radiation pressure counteracts the collapse.

Causal physics cannot affect modes that are so large that they are outside the horizon, thus the quantum fluctuations set up during inflation could only have evolved for modes inside the horizon. The size of the horizon grows as the Universe ages, and more and more modes would therefore have fallen inside the horizon and been affected by causal physics. Once a mode entered the horizon, fluctuations on this scale started collapsing, however the radiation pressure would eventually have caused the mode to bounce back. The battle between gravitational collapse and radiation pressure generates oscillations, resembling the harmonic oscillations of a loaded spring. These oscillations are known as acoustic oscillations, and their imprint are clearly visible in figure 3.10 which shows the power spectrum published by the *Planck* Collaboration. Modes that were frozen out at the peaks and troughs of their oscillations correspond to large fluctuations, since $C_\ell \propto T^2$, and therefore give peaks in the power spectrum, such as those at $\sim 1^\circ$ and $\sim 0.4^\circ$ for the modes that froze out when being maximally compressed the first time and when having gone through one full oscillation, respectively. In this way the odd-numbered peaks correspond to modes that were frozen when maximally compressed and conversely the even-numbered peaks correspond to modes that were frozen when minimally compressed. Modes that are so large that they never entered the horizon were frozen in the initial conditions, and this flat area is called the Sachs-Wolfe plateau.

The red points and error bars in figure 3.10 correspond to measurements by the *Planck* satellite, and the green curve is the power spectrum produced in a universe given by the best-fit cosmological ΛCDM model, rather than a fit to the points. The pale green area shows the predicted variations of this standard model, called cosmic variance, which depends on the number of m-modes for a multipole, which is $2\ell + 1$, and therefore higher for lower multipoles. Changing the parameters of the cosmological model will change

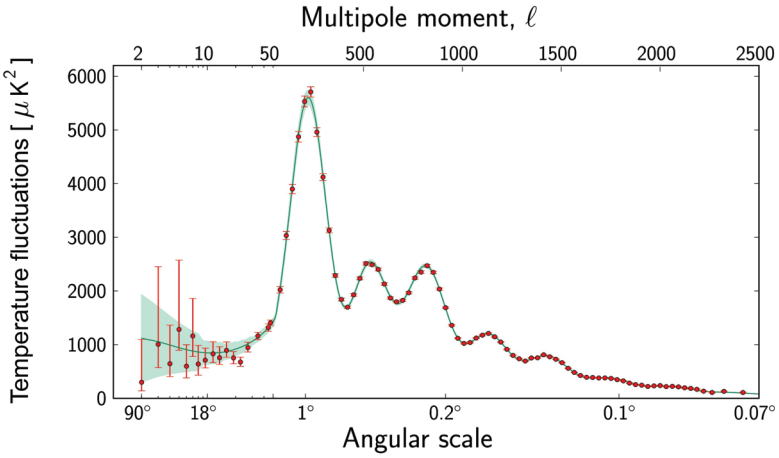


Figure 3.10: The *Planck* power spectrum showing the temperature fluctuations of the CMB radiation. The measurements, with errors, are represented by the red points and green curve is the power spectrum from the best-fit parameters, with the pale area showing the predicted variations of the best-fit standard model. Credit: ESA and the *Planck* Collaboration.

the shape of the power spectrum, for example adding more baryons to the Universe will suppress the even-numbered peaks as compared to the odd-numbered peaks, and increasing/decreasing the curvature of the Universe shifts the acoustic peak structure left/right. In this case the model clearly fits the measurements to high precision for small scales down to 0.07° and shows seven acoustic peaks. However, the anisotropies at the largest scales, of more than a few degrees, show a lack of power in the measurements as compared to the model which we'll come back to in chapter 6.

Part II

Scientific contributions to papers

Introduction

As the title implies, part II contains the candidate's scientific contributions to the papers included in this thesis, and the order of the papers merits an explanation. Instead of following a chronological order from the first to the last project the papers have been ordered to follow a natural time line from observations via data processing to data analysis. This is nearly the opposite order from how the work was performed, where the candidate starting "easy" and went deeper as more knowledge was obtained. The first project consisted of writing a parameter estimation sampler from scratch (chapter 8) and comparing the results to those by the favored sampler. The parameter estimation algorithms used here rely on a likelihood evaluation algorithm, which was the key aspect of the subsequent project in chapter 7 where the candidate dove deeper into the analysis by altering the method of likelihood evaluation. Chapter 6 comprises of the analysis of a statistical phenomenon observed in component separated CMB maps, that goes against the Cosmological Principle, whereas chapter 5 describes a quick and relatively easy component separation technique to produce foreground-reduced maps. The first step in any analysis process is performing the observations, however the observations discussed here are not of the CMB field all the previously mentioned analyses are based on, but of a by product of *Planck*'s CMB analysis, namely clusters of galaxies, which were found through the SZ effect. The candidate has partaken in several spectroscopic follow up runs to confirm or reject candidate clusters.

Chapter 4

Paper I: *Planck* 2013 results. XXIX. *Planck* Catalogue of Sunyaev-Zeldovich Sources

The candidate is currently not listed as an author on the public version of the paper, however the author list will be amended before being published in Astronomy & Astrophysics.

Clusters of galaxies are large scale structures consisting of 100s to 1000s of gravitationally bound galaxies, with a diameter of a few Mpc, and are useful probes to study both dark matter and dark energy. Locating clusters can be done with optical, infrared or X-ray imaging, or through the Sunyaev-Zeldovich (SZ) effect [47, 48]. The SZ effect arises as CMB photons passing through clusters are scattered off hot gas in the cluster, which distorts the shape of the spectrum and the existence of clusters can be inferred through this distortion.

The SZ cluster candidates inferred from *Planck* have to be confirmed through follow-up observations, and for the March 2013 release 276 candidates had gone through observations in the follow-up programme. The Ph.D-candidate contributed to three follow-up runs for this release. *Planck*'s SZ catalogue contains 1227 entries with 861 confirmed clusters, of which 178 are new clusters discovered and confirmed by the *Planck* Collaboration. The remaining 366 candidates will go through follow-up observations when time is granted to the observing programmes, and the Ph.D-candidate has partaken in two post-2013 release follow-up runs.

4.1 Spectroscopic follow-up observations

Certain selection criterion (see section 5 of the paper) were applied to decide which confirmation programme the cluster candidates would go through, based on existing cluster catalogues. The candidates sent to the spectro-

scopic confirmation programme were usually those with a visible galaxy population or optical counter part. Several telescopes of different sizes, ranging from 1 m to 10 m, were used for spectroscopic follow-up, and optimization was applied when assigning clusters to the different telescopes in order to prioritize cluster confirmation and securing the largest number of robust redshifts. The Ph.D-candidate performed spectroscopic follow-up observations in both the Northern and Southern hemispheres at the Nordic Optical Telescope (NOT) on La Palma (Spain) and the New Technology Telescope (NTT) at La Silla (Chile), respectively. Details regarding the telescopes and observations can be found in sections 5.2.1 and 5.3.2 of the paper, however an example of the observations process is given below for the Nordic Optical Telescope.

4.2 Observational procedure at NOT

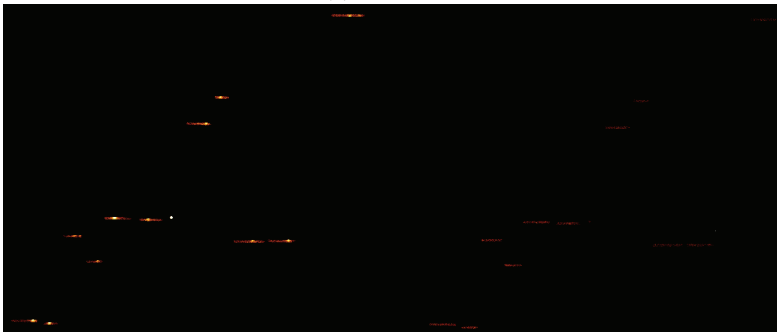
Spectroscopic observations at NOT were performed with the Andalucia Faint Object Spectrograph and Camera (ALFOSC) which has a field of view of $6.4 \times 6.4\,\mathrm{arcmin}^2$. For each candidate cluster a pre-made mask with slits for 10 to 15 galaxies was used to mask out all other sources, except for three small holes corresponding to stars that were utalized for alignment purposes. This mask was created from a pre-image, such as that in figure 4.1b showing the layout of the positions of 12 slits (red rectangles) for galaxies in target cluster 389 and the 3 holes for alignment stars (yellow squares). Furthermore, each mask contains two sets of slits enabling measurements to be taken at two different positions on the CCD chip and the "blank" set of slits can be used to calibrate the background, both of which are very useful for image processing. The left and right sets will be called A and B, respectively. Grism No. 5 was used when taking spectroscopic measurements, which covers a wavelength range of $5000\,\text{Å} - 10\,250\,\text{Å}$.

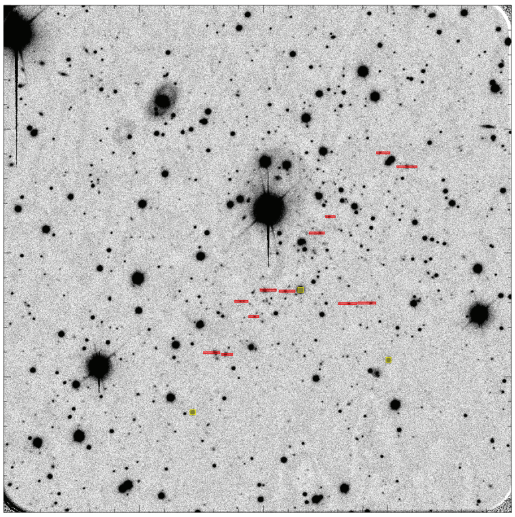
The aim for each target cluster was to take a total of four spectral exposures of between 20 and 30 minutes each, where more distant galaxies require longer exposure time as their galaxies are fainter. These four spectra were split evenly between the two slit sets. Furthermore, between two cluster spectra a set of spectroscopic calibration exposures were taken, using different lamps with the mask and grism in place, according to:

- $3 \times 10\,\mathrm{s}$ exposure with a Halogen lamp, and the calibration lens[1].
- 3 s exposure with a He lamp.
- 3 s exposure with a Ne lamp.
- 30 s exposure with a ThAr lamp.

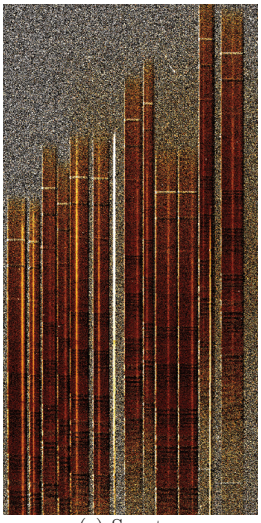
[1] The calibration lens spreads the light from the lamp onto the full CCD.

(a) Alignment





(b) Slits



(c) Spectra

Figure 4.1: Observations of target cluster 389 at NOT. a) Alignment of the cluster galaxies (bright dots) to their corresponding slits in slit set A, where slit set B shows only background. b) Field of view, showing the cluster and other astrophysical foreground or background objects, together with the layout for the mask with slits for 12 galaxies (red rectangles), and 3 holes for alignment stars (yellow squares). c) Spectrum of target 389 with the galaxies centered in slit set A and where the background from slit set B has been subtracted. Courtesy of Håkon Dahle.

Additionally, the calibration spectra of a standard star was taken at least once during the observing run, and bias/dark frames were taken daily during twilight. All these factors make the post-processing more accurate.

The mask holder on ALFOSC can fit 3 masks at a time, and the orientation of these masks were checked with pre-imaging before or during twilight each evening. Towards the end of astronomical twilight it is dark enough to start performing observations and the first step after slewing to a target cluster is to check the orientation of the field of view and adjust if necessary. Further, the alignment of the target galaxies on the slits was acheived using ALFOSC target acquisition by running `mosstrong` which uses images of the mask and of the field of the candidate cluster. By selecting the three star holes on the mask image and the corresponding stars on the field of view image, the required telescope offset is computed and the `mosmove` command will make the necessary adjustments to ensure that the target galaxies appear inside the slits. Minor tweaking was performed directly through the telescope offset (`teloffset`) to ensure that the galaxies were well centered on their corresponding slits. The alignment of the galaxies, represented by bright dots, to their corresponding slits in slit set A is shown in figure 4.1a for target 389, together with the empty slits of slit set B.

Once as many galaxies as possible[2] have been placed on the slits the spectra can be captured, followed by the first set of lamp calibrations. The order of observations usually follow the ABBA rule, where the first spectra and calibration occur with the galaxies on the A slits, followed by two sets of spectra and calibrations with the galaxies on the B slits, and then another spectra at the A slits. The approximate `teloffset` required to dither from slit set A to B can be calculated from the mask image by computing the average number of pixels to get from set A to set B based on the holes/slits, where each pixel corresponds to 0.19 arcsec. After dithering between slit sets the alignment of galaxies on the slits has to be checked and adjusted where appropriate before continuing the spectral observations. Figure 4.1c shows the spectra obtained when placing the galaxies on the slits of set A, where the background spectra from the slits in set B has been subtracted. In other words the spectra in the slits have been background-cleaned, which is the first step in the post-processing stage to compute the redshift to the cluster in question.

[2]In certain cases the mask quality or misalignment of some slits result in a scenario where some galaxies fall outside their intended slit and the observers have to make a choice on which galaxies are most/least important. In these cases it is prudent to capture the spectra for as many galaxies as possible.

Chapter 5

Paper II: *Planck* 2013 results. XII. Component Separation

At the end of the day it is the "pure" CMB signal that cosmologists crave for and a lot of work has to be done to disentangle this component from the foregrounds and noise (equation (3.5)), through component separation techniques. There are a variety of different ways to perform this separation, and choices have to be made between assuming one knows nothing or knowing something, and which domain to perform the analyses in, which can be in either pixel or harmonic domain. The *Planck* Collaboration has implemented four different analyses through combinations of these choices, as can be seen in table 5.1. These four component separation algorithms are called `Commander-Ruler`, `NILC`, `SEVEM` and `SMICA`. The analyses that assume to know nothing aim to minimize the variance of the CMB signal (minimizing the error in the result), whereas those assuming to know something involve fitting models of the CMB radiation and foregrounds. Having several component separation codes is an advantage as they can be used to check consistency by comparing and contrasting the maps themselves and results from analyses run on these maps.

	Know something	**Know nothing**
Pixel domain	`Commander-Ruler`	`SEVEM`
Not pixel domain	`SMICA`	`NILC`

Table 5.1: Classification of the component separation techniques by the *Planck* Collaboration into the state of knowledge assumption and analysis domain.

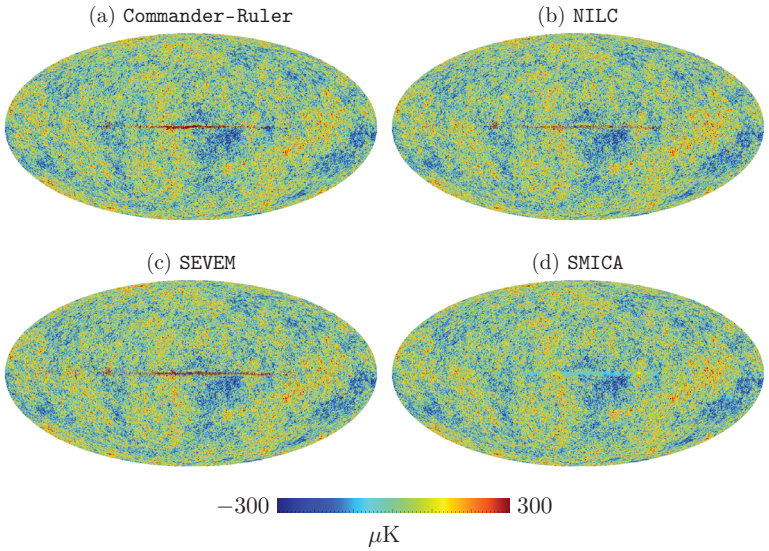


Figure 5.1: Component separated CMB maps for the four component separation techniques by the *Planck* Collaboration; a) `Commander-Ruler`, b) `NILC`, c) `SEVEM`, and d) `SMICA`. Note that the `SMICA` map has been filled in smoothly inside a 3% mask, whereas the other three solutions show their galactic residual.

The component separated maps resulting from the four techniques are shown in figure 5.1 and overall they look very similar, from which one can conclude that the solutions are in good agreement. However, the radiation coming from the Milky Way galaxy is very strong and all four component separation techniques leave a band along the galactic plane, which are quite dissimilar for the different codes. Note that for the `SMICA` solution (figure 5.1d) a 3% mask has been applied to remove the left-over galaxy contamination and some strong sources at higher latitudes, and this area has subsequently been smoothly filled in.

By taking the difference between pairs of CMB maps the dissimilarity between the solutions can be investigated, as shown in figure 5.2 where darker red and blue indicate greater differences between the solutions, with a cut-off at $\pm 30\,\mu$K. As already mentioned the solutions are visibly different along the galactic plan, but clearly the solutions also vary at higher latitudes at a level of $\pm 5\,\mu$K. The most noticeable of these high-latitude differences is a negative signal along the ecliptic plane that is present in all figures involving `Commander-Ruler` (figures 5.2a, 5.2b and 5.2c). The zodiacal light emission (ZLE) [69, 70] lies in an S-shaped band around the ecliptic plane and could

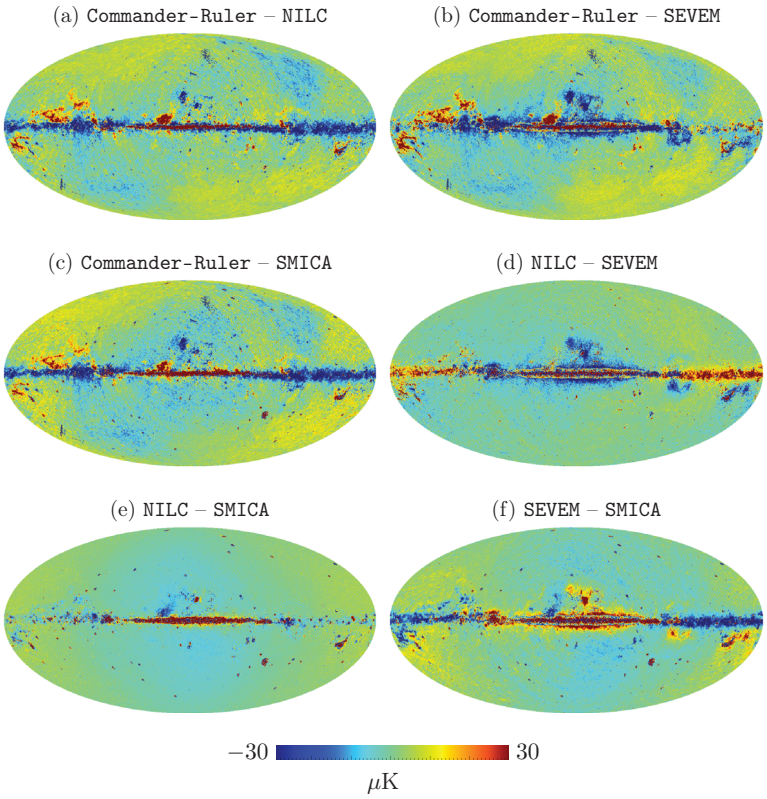


Figure 5.2: Difference maps for pairs of component separated CMB maps from `Commander-Ruler`, `NILC`, `SEVEM` and `SMICA` (figure 5.1) that have been downgraded to $N_{\text{side}} = 128$ to emphasize the large scale differences. The captions indicate the subtraction order, which is alphabetical.

well be the cause of the difference between the `Commander-Ruler` solution and the others. The ZLE is a high frequency signal caused by the reflection of sunlight off dust in the Solar system, however its spectrum is exceedingly close to that of galactic thermal dust which makes it very difficult to tell these two foregrounds apart. The differences in the solutions can either be due to the lack of high frequency information in the `Commander-Ruler` analysis[1] making it unable to distinguish the ZLE from thermal dust, or due to ZLE leakage from the high frequency channels into the CMB solutions for the `NILC`, `SEVEM` and `SMICA` analyses. A more detailed ZLE model is

[1]The `Commander-Ruler` framework excludes the two highest channels (545 GHz and 857 GHz) from the analysis.

required before its role in this dissimilarity can be solved, however a third possibility is of course that the ZLE is not responsible for the discrepancy at all. Regardless of where the difference comes from its impact on power spectrum and consequently on parameter estimation should be small, although it may be important for studies of large scale phenomenon [71]. Such studies should therefore be performed using all four component separation solutions to check the consistency of the results. See the *Planck* 2013 release paper on component separation [72] and references therein for more details on the methods and comparisons of results.

5.1 Power spectrum model at large scales

The candidate's contribution to the paper is found in section 6.2.2 and comprises of low-ℓ power spectrum estimation for low resolution versions of the component separated CMB maps. Given an input CMB-map and *Planck*'s fiducial power spectrum, $C_\ell^{\rm fid}$, the best-fit power spectrum parameters for a simple model, given by

$$C_\ell = q \left(\frac{\ell}{\ell_{\rm piv}} \right)^n C_\ell^{\rm fid}, \tag{5.1}$$

can be found, where q and n are the parameters for the power spectrum amplitude and tilt, respectively, and $\ell_{\rm piv}$ is the pivot multipole for the correction.

Equivalent low resolution maps from the four component separation techniques were used for the analyses. These maps were subsequently smoothed with a 6° Gaussian beam and downgraded to $N_{\rm side} = 32$ to capture only the large scale anisotropies and reduce computational cost associated with analysis of higher resolution maps. A simple grid-based likelihood analysis was performed to investigate the two-dimensional parameter space[2], and figure 5.3 shows the resulting contours for `Commander-Ruler`, `NILC`, `SEVEM` and `SMICA` in green, blue, red and orange curves, respectively. The innermost contour represents the area close the maximum likelihood and the outer three contours are those for 1σ, 2σ and 3σ. The black $\times$ represents the fiducial $(q, n) = (1, 0)$ model, and there is clearly tension between this model and the best-fit results from the low-ℓ analysis, although only at a 1.7σ level.

The four solutions are in excellent agreement with one another with the differences between their solutions being less than 0.1σ, which shows that the ecliptic band difference discussed above has no impact on these results. This tension has also been investigated using the full *Planck* likelihood [73, 19] where the significances for the low-ℓ power spectrum discrepancy was found to be higher.

[2]This analysis was performed using the dipole modulation code that will be presented in the next section.

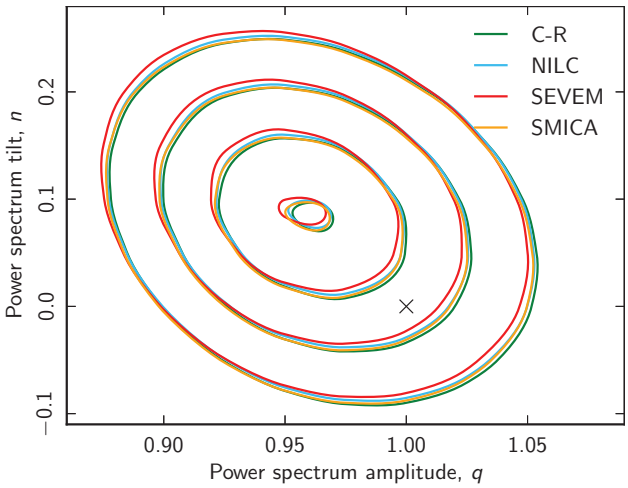


Figure 5.3: Contours for the power spectrum amplitude and tilt parameters for `Commander`, `NILC`, `SEVEM` and `SMICA` maps smoothed to 6° in, respectively, green, blue, red and orange. The three outermost contours correspond to the 1σ, 2σ and 3σ bounds, the innermost contour represents the area close to the best-fit values and the $\times$ is located at the fiducial power spectrum parameters $(q, n) = (1, 0)$.

5.2 Internal Linear combination

The work in this section did not make it into the paper, but it will nonetheless be discussed as part of scientific contributions to the paper as it provides an alternative approach to foreground-reduction for *Planck*. The following was primarily based on work by Eriksen et al. [74] that was previously used for *WMAP* analysis.

A much quicker method for producing a foreground-cleaned map than the component separation techniques discussed above is by combining the sky maps from several frequency channels through internal linear combination (ILC). However, this method is also much less accurate than the full component separation techniques and care should be taken when using such maps for scientific analysis.

The CMB temperature signal is independent of the measurement frequency, ν_i, thus the difference in measured temperature between the frequency maps comes solely from the foreground contributions, $T_{\rm fg}$. The temperature of the foregrounds are frequency dependent, meaning that the temperature field across the sky for frequency channel i can be written as

$$T(\nu_i) = T_{\rm CMB} + T_{\rm fg}(\nu_i), \tag{5.2}$$

where $T_{\rm CMB}$ is the underlying CMB temperature field.

The aim of the ILC method is to produce a temperature field, T, which minimizes the contribution from the foreground contaminants while retaining the CMB temperature field, which can be achieved through the linear combination of the measured temperature field for k frequency bands,

$$T = \sum_{i=1}^{k} w_i T(\nu_i), \tag{5.3}$$

where w_i is the weight assigned to the temperature field for measurements at frequency ν_i. The requirement that $\sum_{i=1}^{k} w_i = 1$ is imposed to ensure that the CMB temperature is not weighted by an unwanted factor. The remaining temperature field can be written as,

$$T = T_{\rm CMB} + \sum_{i=1}^{k} w_i T_{\rm fg}(\nu_i), \tag{5.4}$$

where we have used that $\sum_{i=1}^{k} w_i T_{\rm CMB} = T_{\rm CMB}$ since the CMB signal is independent of frequency and the weights are required to add to 1. The weights are selected to minimize the impact of the foreground signal and the most convenient way to compute them is such that the variance of T is minimized, which is given by

$$\mathrm{var}(T) = \mathrm{var}(T_{\rm CMB}) + \mathrm{var}(\sum_{i=1}^{k} w_i T_{\rm fg}(\nu_i)), \tag{5.5}$$

since the temperature component of the CMB signal is statistically independent from that of the foreground and noise. The variance of T can be shown [74] to simplify to,

$$\mathrm{var}(T) = \langle T^2 \rangle - \langle T \rangle^2 = \sum_{i,j=1}^{k} w_i C_{ij} w_j, \qquad \text{where} \tag{5.6}$$

$$C_{ij} = \langle \Delta T_i \Delta T_j \rangle = \frac{1}{N_{\rm pix}} \sum_{p=1}^{N_{\rm pix}} (T^i(p) - \bar{T}^i)(T^j(p) - \bar{T}^j) \tag{5.7}$$

is the map-to-map covariance matrix. The computation of the weights required for each channel in order to minimize the variance is thereby reduced to minimizing

$$f(\boldsymbol{w}) = \sum_{i,j=1}^{k} w_i C_{ij} w_j, \tag{5.8}$$

under the condition,

$$g(\boldsymbol{w}) = \sum_{i=1}^{k} w_i = 1, \tag{5.9}$$

which in the case of Lagrange multipliers means looking for a solution amongst the points $\boldsymbol{w}_0$ that satisfy

$$\Delta f(\boldsymbol{w}_0) = \lambda \Delta g(\boldsymbol{w}_0), \tag{5.10}$$

where λ is an arbitrary constant. In other words, using Lagrange multipliers to minimize the foreground contribution requires one to locate points for which the gradients of f and g are parallel. Solving this equation leads to weights given by,

$$w_i = \frac{\sum_{j=1}^{k} C_{ij}^{-1}}{\sum_{j,l=1}^{k} C_{jl}^{-1}}. \tag{5.11}$$

Armed with equation (5.11) to calculate the weights for each of the k frequency maps, the maps can be combined using equation (5.3) to give a foreground-reduced ILC map of the full sky. However, as we saw in section 3.3, the properties of the foregrounds change across the sky, and in particular tend to be more prominent around the galactic plane than at the poles. Therefore the sky is split into separate regions to account for variations in foreground intensities and map weights are computed for each of these regions separately. The map of the regions used to compute weights for the ILC map produced by the *WMAP* team, shown in figure 5.4, has 12 separate regions. 11 of these regions cover different patches in and around the galactic plane and the last, in dark blue, covers the remaining area at higher latitudes and some patches in the galactic plane. Using several regions results in a set of weights for each region. Each of these sets are used to combine the frequency channels in their patch, thereby producing a foreground-reduced patch for that region. These patched foreground-reduced maps are subsequently recombined to create the full map, with a Gaussian $90'$ FWHM weighting applied at the edges of the regions to prevent ugly discontinuities.

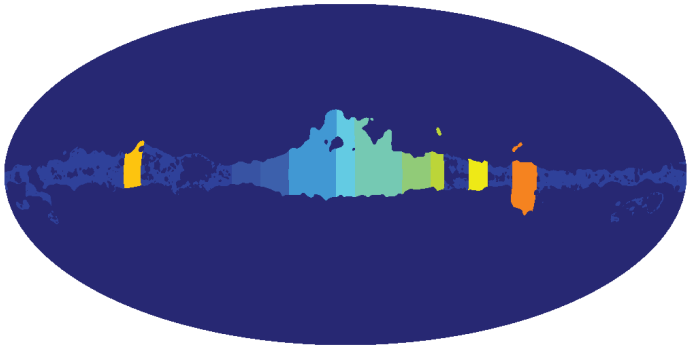


Figure 5.4: Map of the 12 regions used for the ILC map construction for *WMAP*'s 9-year release where the colors represent individual regions - available from LAMBDA (`http://lambda.gsfc.nasa.gov/`).

5.3 ILC map using *WMAP* data

The *WMAP* team has produced foreground-reduced ILC maps to accompany the data releases, like the 9-year ILC map shown in figure 5.5a[3]. The official ILC maps are created through a linear combination of the 1° smoothed coadded I maps for the 5 frequency bands[4], where the weights assigned to the different frequency maps are computed through a nonlinear search to minimize var(T) [75] for the 12 regions shown in figure 5.4. The resulting patches are recombined in the same manner as that for the Lagrange multiplier method described above.

Figure 5.5b shows the foreground-reduced ILC map produced by the Lagrange multiplier method, where the same input maps as for the official *WMAP* ILC method were used. Both of these ILC maps look reasonably well foreground cleaned, with only a hint of remnant foregrounds in the central region of the galactic plane. The weights the two methods compute for the 12 regions are shown in table 5.2 together with the average deviation per region. It is clear that the weights assigned to the 5 frequency bands differ for all map regions. This difference naturally leads to a difference in the maps the two methods produce, however the difference is not noticeable by eye. Furthermore, by subtracting the ILC map produced by the official *WMAP* solution from that resulting from the Lagrange multiplier method, the difference in the two maps becomes clear, as shown in figure 5.5c. The strongest deviations of around $\pm 50\,\mu$K lie in the galactic plane region, but there are also degree sized deviations at a $\pm 20\,\mu$K level at higher latitudes.

[3]Obtained from LAMBDA (`http://lambda.gsfc.nasa.gov/`)

[4]The smoothed I maps per frequency band for the 9-year data release are available at `http://lambda.gsfc.nasa.gov/product/map/dr5/maps_band_smth_r9_i_9yr_get.cfm`.

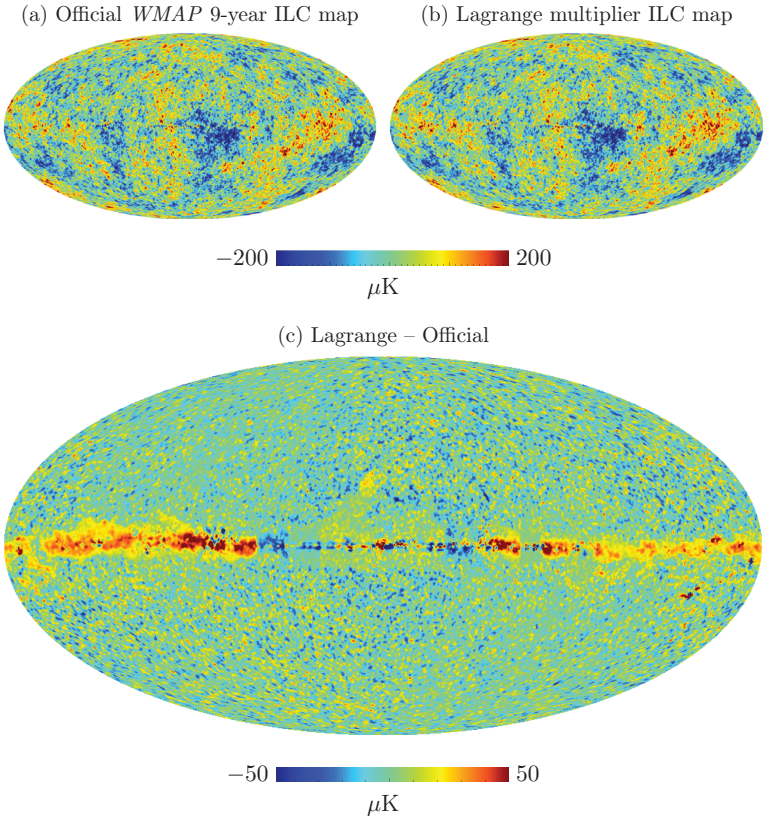


Figure 5.5: Internal linear combination maps produced by a) the official non-linear *WMAP* method, and b) the Lagrange multiplier method, using the smoothed coadded I maps for the 5 frequency bands of the *WMAP* 9-year release, and c) shows the difference obtained when subtracting a) from b).

	***WMAP*9 non-linear**					**Lagrange multipliers**					**Deviation**
	K	Ka	Q	V	W	K	Ka	Q	V	W	
Region 0	0.1555	-0.7572	-0.2689	2.2845	-0.4138	0.0826	-0.7645	0.2759	1.4676	-0.0616	0.36
Region 1	0.0375	-0.5137	0.0223	2.0378	-0.5839	0.0783	-0.6762	0.1690	1.8670	-0.4381	0.13
Region 2	0.0325	-0.3585	-0.3103	1.8521	-0.2157	0.0122	-0.2759	-0.3716	1.8587	-0.2234	0.04
Region 3	-0.0910	0.1741	-0.6267	1.5870	-0.0433	-0.0853	0.1299	-0.5682	1.5606	-0.0371	0.03
Region 4	-0.0762	0.0907	-0.4273	0.9707	0.4421	-0.0572	0.0030	-0.3495	0.9545	0.4492	0.04
Region 5	0.1998	-0.7758	-0.4295	2.4684	-0.4629	0.2598	-1.0648	-0.1789	2.4834	-0.4995	0.13
Region 6	-0.0880	0.1712	-0.5306	1.0097	0.4378	-0.1042	0.2210	-0.5373	0.9566	0.4640	0.03
Region 7	0.1578	-0.8074	-0.0923	2.1966	-0.4547	0.0620	-0.4345	-0.2977	1.9495	-0.2792	0.22
Region 8	0.1992	-0.1736	-1.8081	3.7271	-0.9446	0.2366	-0.6198	-1.0205	3.1678	-0.7641	0.40
Region 9	-0.0813	-0.1579	-0.0551	1.2108	0.0836	-0.1309	-0.0134	-0.0672	1.0870	0.1244	0.07
Region 10	0.1717	-0.8713	-0.1700	2.8314	-0.9618	0.1760	-0.8479	-0.2286	2.7921	-0.8916	0.04
Region 11	0.2353	-0.8325	-0.6333	2.8603	-0.6298	0.2370	-0.8587	-0.5758	2.7942	-0.5967	0.04

Table 5.2: Left: Official *WMAP* 9-year ILC weights from Bennett et al. [33] using non-linear search method; Middle: ILC weights by Lagrange multiplier method; Right: Average deviation between official and Lagrange weights.

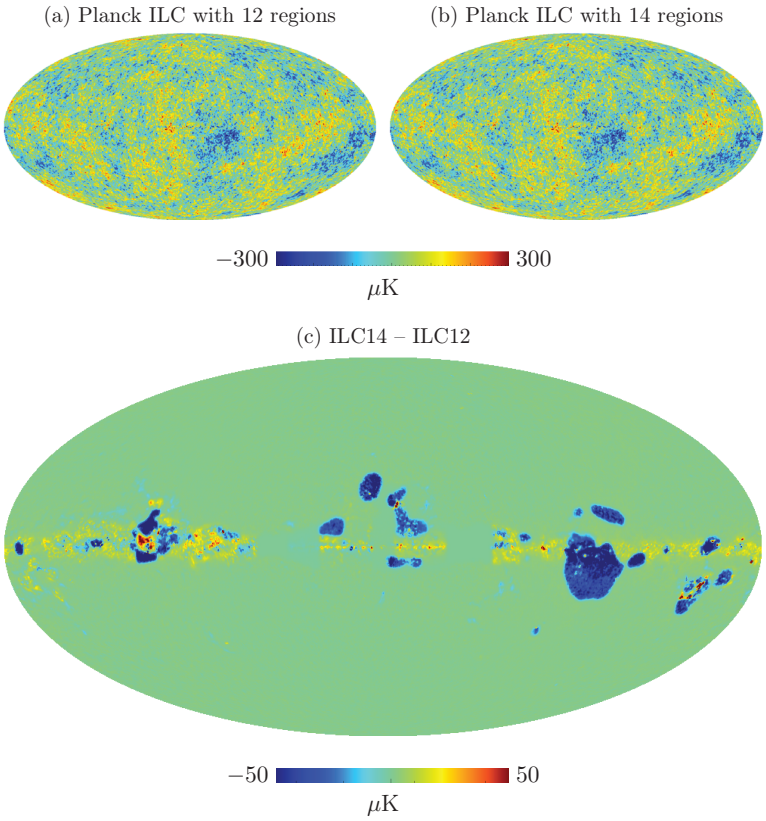


Figure 5.6: Foreground-reduced ILC maps for the seven lowest *Planck* frequencies produced using Lagrange multipliers with a) 12 regions, and b) 14 regions, and where c) shows the difference between the two solutions.

5.4 ILC map using *Planck* data

The *Planck* satellite made observations at nine frequency bands, and these can be combined using internal linear combination to produce a foreground-cleaned map of the CMB. There is no official ILC map produced by the *Planck* Collaboration, but figure 5.6a shows an unofficial map made by using Lagrange multipliers and the region definitions as for *WMAP*, with the seven lowest frequency *Planck* maps[5] smoothed to $40'$.

[5]The frequency maps for *Planck* are available on *Planck* Legacy Archive (`http://www.sciops.esa.int/index.php?page=Planck_Legacy_Archive&project=planck`)

Figure 5.7: Extended region map for *Planck* created from the original *WMAP* 9-year region file by adding two additional regions for the free-free (dark orange) and CO (red) emission areas based on *Planck* maps of these emissions.

The *Planck* Collaboration's Gibbs sampling based component separation routine `Commander`, does not only produce the component separated CMB map, but also maps of the different foregrounds. Of particular interest are the CO and free-free foreground maps, which can be used to create two additional regions in the ILC region map to represent the areas where these two foregrounds are dominant. The area of the extra region is selected by finding the areas with emissions above a certain threshold that have at least half the neighbors inside a circle of radius 2° that also have values above the threshold. The extended ILC region map is shown in figure 5.7, where the additional regions to capture the emission areas of free-free and CO are colored dark orange and red, respectively. This region map was used with the same seven *Planck* frequency maps as for the analysis using 12 regions, and the resulting ILC map is shown in figure 5.6b.

By eye these maps look indistinguishable, but as figure 5.6c shows the additional regions do create differences between the two solutions of as much as $50\,\mu$K. However, comparing the two ILC maps directly does not really give an indication of which is a better fit to the real CMB sky, but such a distinction can be made by comparing them to the *Planck* Collaboration's component separated maps. The ILC maps are smoothed with a $40'$ Gaussian beam and contain the monopol and dipole terms, thus before a comparison can be made the $\ell = 0$ and $\ell = 1$ terms are removed from the ILC maps, and the component separated maps are smoothed with a $40'$ Gaussian beam. Furthermore, the maps are also downgraded to $N_{\mathrm{side}} = 128$, before producing the difference maps shown in figure 5.8. The maps from top to bottom show the `Commander-Ruler`, `NILC`, `SEVEM` and `SMICA` solutions subtracted from

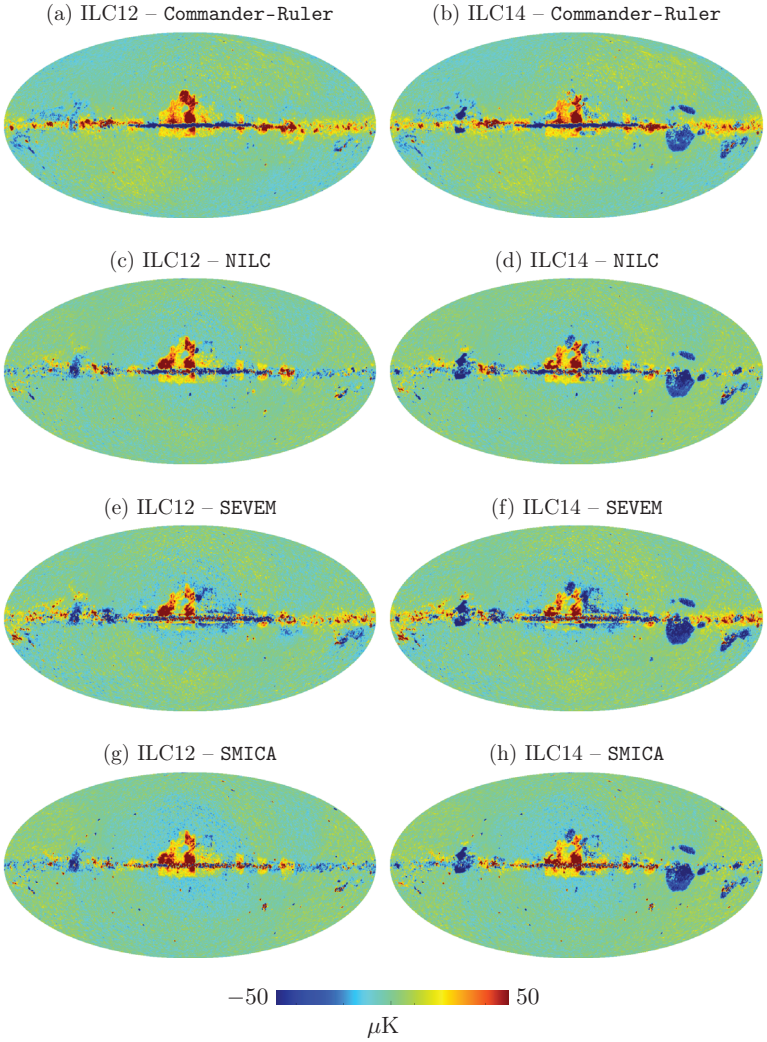


Figure 5.8: Difference maps produced by, from top to bottom, subtracting the `Commander-Ruler`, `NILC`, `SEVEM` and `SMICA` solutions from the ILC maps, where maps on the left-hand side are those using the ILC map produced with the 12 *WMAP* regions, and the right-hand side are those for 14 regions.

the ILC map, where the maps on the left-hand side are those for the ILC solution using the 12 *WMAP* regions, and the right-hand side for that using 14 regions. Rather surprisingly, the addition of two extra regions to account for the free-free and CO emission areas appear to make the ILC solution less accurate, as compared to the component separated maps. The most noticeable difference is the large negative spots above and below the galactic plane on the right-hand side, which are only visible for the 14 region ILC map. This area corresponds to a region with free-free emission, and upon closer inspection most of the areas this region covers appear to have a lack of power in the ILC solution with 14 regions. On the other hand, the regions corresponding to the CO emission appear to be slightly better foreground-reduced in the 14 region ILC map than that for 12 regions, although these differences are very small and therefore quite difficult to see.

5.5 Way forward

The current ILC analysis based on Lagrange multipliers was limited to *Planck*'s nominal data set of 15.5 months, but as mentioned in section 3.2.2 the HFI and LFI instruments collected data for a total of 30 and 48 months, respectively. The most obvious continuation of this project would therefore be to perform the analyses on the full data set. Furthermore, the common smoothing applied to the frequency maps using a FWHM= $40'$ Gaussian beam was implemented by using approximated Gaussian shaped input beams for the different frequency maps. This should be extended to utilize the beam shapes produced by the *Planck* Collaboration to increase the accuracy of the applied smoothing.

Additionally, the analysis was performed with only the seven lowest frequency channels and a more complete analysis should be performed where all nine frequency maps are included, or to analysis limited to different subsets of eight, seven or fewer maps. Comparisons of the results from these analysis will be informative as to the contribution the different frequency maps make to the analysis, and which bands contribute foreground residuals.

Chapter 6

Paper III: *Planck* 2013 results. XXIII. Isotropy and Statistics of the CMB

The Cosmological Principle, which is one of the fundamental assumptions CMB analysis relies on, states that the Universe is homogeneous and isotropic, and therefore it follows that also the CMB sky should be homogeneous and isotropic. However, already during the analysis of the first year *WMAP* data [76] a number of anisotropies were detected on the largest scales - an anomalously large cold spot in the Southern Hemisphere (known as the Cold Spot) [77, 78], the strong alignment of the quadrupole and octopole axes [79, 80] and the asymmetry in the power fluctuations on opposite hemisphere [81, 82, 83]. These asymmetries have been investigated for all the subsequent *WMAP* data releases (eg. [84, 85, 86, 87]) and found to persist.

Paper XXIII [71] is dedicated to the isotropy and statistics of the CMB radiation as seen by *Planck*, and includes analyses of the anisotropies mentioned above, as well as several other anisotropies. This section will, however, only concentrate on the contribution by the candidate, entailing the analysis of the hemispherical power asymmetry at large scales using a dipole modulation model as discussed in section 5.5.2 of the paper.

The hemispherical power analysis of the first year release from *WMAP* showed that there was an asymmetry between the power in opposite hemispheres, with an excess power found in the southern hemisphere. For later data releases the analysis was extended to a modulation modeled as a dipolar modulation field, first introduced by Gordon et al. [88], which thereby gives a best-fit direction and amplitude for the asymmetry. The power asymmetry analysis presented here is a continuation of such analyses, by Eriksen et al. [89] and Hoftuft et al. [90], which due to the computational cost of analyzing high resolution data were restricted to the very largest scales with $\ell_{\rm max} = 64$. Other teams have looked at the asymmetry in the *WMAP* data on smaller

scales, to $\ell_{\rm max} = 600$ [91, 87], which is discussed in section 5.5.1 for the *Planck* release. The analyses based on *WMAP* data have produced results that are in good agreement, particularly with regards to the direction of the power asymmetry.

The construction and operation of the *Planck* and *WMAP* satellites were very different, for example with regards to the instrument technology used, frequency band choices and scanning strategy of the sky, which means that they suffer from different systematics. Additionally the component separation methods implemented by the *Planck* Collaboration clean foregrounds very well, as every known foreground has been taken into account. Therefore, if the anisotropies are still present in the analysis of the *Planck* data and in agreement with results from *WMAP* then they are unlikely to be due to systematic effects or foreground contamination. Furthermore, all of these anomalies put the cosmological standard model under pressure since the model has no explanation for their existence, and they could therefore be the indication of new physics beyond the standard cosmological ΛCDM model.

6.1 Dipole modulation model

In the framework introduced by Gordon et al. [88], the CMB signal, $s(\hat{\boldsymbol{n}})$, is modeled as an isotropic and Gaussian random field with power spectrum C_ℓ. This field is modulated by a dipolar modulation field given by

$$f(\hat{\boldsymbol{n}}) = A(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{p}}), \tag{6.1}$$

where A and $\hat{\boldsymbol{p}}$ are the amplitude and preferred direction of the modulation field, respectively. By taking into account the instrumental noise, $n(\hat{\boldsymbol{n}})$, the observed data can be written as,

$$d(\hat{\boldsymbol{n}}) = [1 + f(\hat{\boldsymbol{n}})]s(\hat{\boldsymbol{n}}) + n(\hat{\boldsymbol{n}}). \tag{6.2}$$

The underlying low-ℓ power spectrum of the CMB is computed from *Planck*'s fiducial power spectrum, $C_\ell^{\rm fid}$, by

$$C_\ell = q \left(\frac{\ell}{\ell_{\rm piv}} \right)^n C_\ell^{\rm fid}, \tag{6.3}$$

where q and n are the power spectrum amplitude and tilt parameters, which are applied around the pivot multipole of $\ell_{\rm piv}$. These two parameters are free parameters and allowed to vary to find their best-fit values.

6.2 Grid-based maximum likelihood search

The previous analyses using this dipolar modulation were implemented using an MCMC method to sample parameter space, which was adapted for the *Planck* analysis to use an adaptive grid-based sampling method instead. Since the dipole modulation parameters, A and $\hat{\boldsymbol{p}}$, and the power spectrum parameters q and n are almost totally uncorrelated, the grid sampling is performed by iterating between finding the maximum likelihood parameters in $A - \hat{\boldsymbol{p}}$ space and in $q - n$ space, while fixing the values of the other parameter set.

The first iteration maps dipole modulation space $(A - \hat{\boldsymbol{p}})$ with the power spectrum fixed to the fiducial power spectrum, and performs a maximum likelihood search to find the best-fit parameters, $(A, \hat{\boldsymbol{p}}) = (A, \hat{\boldsymbol{p}})_{\rm ML}$. The dipole modulation parameters are fixed to the obtained maximum likelihood values for the second iteration, which investigates power spectrum space $(q - n)$ to locate the maximum likelihood parameters, $(q, n) = (q, n)_{\rm ML}$. The third iteration is identical to the first, but with the power spectrum parameters fixed to the best-fit values from the previous iteration. The maximum likelihood parameters resulting from the third iteration may be slightly different from that of the first iteration, but not by a significant amount, which shows the negligible correlation between dipole modulation and power spectrum.

This iteration method can be summarized as

1. fix $(q, n) = (1, 0)$ [1] - find $(A, \hat{\boldsymbol{p}})_{\rm ML}$
2. fix $(A, \hat{\boldsymbol{p}}) = (A, \hat{\boldsymbol{p}})_{\rm ML}$ - find $(q, n)_{\rm ML}$
3. fix $(q, n) = (q, n)_{\rm ML}$ - find $(A, \hat{\boldsymbol{p}})_{\rm ML,2}$

The grid setup for the modulation amplitude, and the power spectrum amplitude and tilt are very straight forward. The start and end points, and the grid sizes are provided by the user, which defines the one-dimensional grid for the parameter in question.

The setup for the modulation direction is however not so straight forward. As mentioned in section 3.3, the `HEALPix` map with the lowest resolution consists of 12 pixels and for the subsequent level each pixels of the previous level are split into 4 new pixels. The investigation of the preferred direction of the dipole modulation is therefore first performed with the possible directions given by the centers of the 12 $N_{\rm side} = 1$ pixels for this low resolution direction case. For the next iteration the possible modulation directions are given by the centers of the $N_{\rm side} = 2$ pixels, of which there are 48. However, if the likelihood of a pixel at the previous $N_{\rm side}$ was below a certain cutoff value the

[1] This means that the power spectrum is set to the fiducial power spectrum, $C_\ell = C_\ell^{\rm fid}$

likelihoods for the 4 subpixels are set to equal the value of the higher level pixel instead of wasting time on their likelihood evaluations. This approach continues up to a maximum $N_{\rm side}$ as set by the user, where the threshold decreases with increasing $N_{\rm side}$, thereby decreasing the area investigated. Setting cutoff likelihoods at each level ensures that only areas that show signatures of something interesting are investigated further. This reduces computational cost since the number of computations for each improvement in resolution is lower than the total number of pixels, and also lower than the number of subpixels with regards to the computations at the previous resolution level. This is illustrated in figure 6.1, where maps a), b) and c) show the pixels for $N_{\rm side} = 1$, 2 and 4, respectively, with color representing the corresponding likelihoods computed at the center of the pixels. The color scaling is chosen such that the pixel with the highest likelihood is deep red and pixels that fall below the threshold are in the darkest blue, which are the pixels that will be excluded for the next iteration. The number of pixels at $N_{\rm side}$ of $1, 2$ and 4 that will be investigated further are, respectively, 4, 7 and 16, although the 4th pixel at $N_{\rm side} = 1$ (right hand side above the plane) is also very dark blue as its likelihood value happens to be very close to the cutoff likelihood. Map d) shows the likelihood values for the final map iteration with $N_{\rm side} = 8$, where the adaptive structure is visible.

To compute the significance of the modulation result obtained for a particular map the log-likelihood difference, $\Delta \ln \mathcal{L}$, is computed, which comes from computing likelihoods with $(\hat{\boldsymbol{p}}) = (\hat{\boldsymbol{p}}_{\rm ML,2})$, $(q, n) = (q, n)_{\rm ML}$ and letting the modulation amplitude equal 0 or $A_{\rm ML,2}$. Thus $\Delta \ln \mathcal{L} = \ln \mathcal{L}(A = A_{\rm ML,2}) - \ln \mathcal{L}(A = 0)$, which can be used to find the significance of a dipole model with $(A, \hat{\boldsymbol{p}}) = (A, \hat{\boldsymbol{p}})_{\rm ML,2}$ as compared to no modulation.

The relationship between log-likelihood difference and chi-square, χ^2, is given by

$$\chi^2 = 2\Delta \ln \mathcal{L}, \tag{6.4}$$

and the significance is computed as

$$\text{significance} = \sqrt{\chi^2}. \tag{6.5}$$

However, the log-likelihood difference above is computed for 3 degrees of freedom (dof) and the significance should be computed for 1 degree of freedom to find the significance of the $A = A_{\rm ML,2}$ result. To compute the correct significance the following procedure is used:

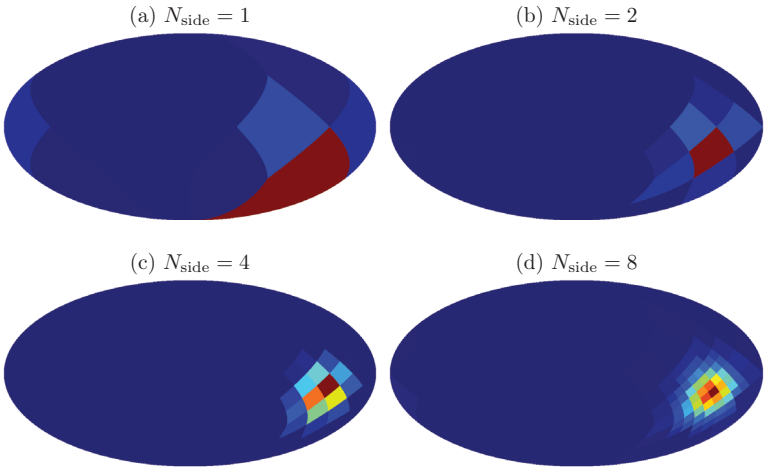


Figure 6.1: Illustration of the adaptive grid method with modulation directions given by the centers of `HEALPix` pixels at a) $N_{\rm side} = 1$, b) $N_{\rm side} = 2$, c) $N_{\rm side} = 4$ and d) $N_{\rm side} = 8$, where only pixels above a certain threshold are investigated for the next iteration. For the first three refinement levels the threshold is represented by the deep blue, and any pixel with another color is investigated further.

1. Compute chi-square from $\Delta \ln \mathcal{L}$: $\chi^2_3 = 2\Delta \ln \mathcal{L}$
2. Compute tail probability for 3 dof from χ^2_3 : $P(\chi^2_3)$
3. Equate tail probabilities for 1 and 3 dof : $P(\chi^2_1) = P(\chi^2_3)$
4. Compute chi-square for 1 dof from $P(\chi^2_1)$: χ^2_1
5. Compute significance from χ^2_1 : Significance$= \sqrt{\chi^2_1}$

For n degrees of freedom $P(\chi^2_n)$, the tail probability, refers to the area under the probability distribution to the right of the chi-square value of χ^2_n, and conversely χ^2_n is the chi-square value for which the tail probability equals $P(\chi^2_n)$. The chi-square and tail probability computations were performed using the Chi Square Calculator by Dr. R. Webster West[2].

[2]Dr. West's Chi-square calculator is available at `http://www.stat.tamu.edu/~west/applets/chisqdemo.html`.

6.3 Results

Analyses of the hemispherical power asymmetry with the dipole modulation model was performed on low resolution $N_{\rm side} = 32$ maps for the 2013 *Planck* release. These were smoothed with Gaussian beams with full width half max (FWHM) of $5^\circ, 6^\circ, 7^\circ, 8^\circ, 9^\circ$, and 10° and subsequently had a noise of $1\,\mu$K added to keep the covariance matrices positive definite. The six different smoothing scales were applied to the low resolution maps for each of the four component separation codes[3]. The resulting maps were processed as explained in section 6.2 with $\ell_{\rm max} = 95$, $\ell_{\rm piv} = 30$ and a highest direction pixelization of $N_{\rm side} = 8$. The numeric results from these twenty-four analyses are found in table 6.1 together with some previously published dipole modulation results from MCMC analyses of ILC maps for the 3 and 5 year releases [89, 90].

6.3.1 Dipole modulation amplitude and direction

The numeric values of the modulation amplitude and direction for the different input maps are quoted in the third and fourth columns of table 6.1, and clearly there is good agreement between the different component separation techniques for a given smoothing scale. The marginals of the amplitude for different component separation techniques with smoothing scale of 5° are shown in figure 6.2a for `Commander`, `NILC`, `SEVEM` and `SMICA` in green, blue, red and orange, respectively. The solid lines are the marginals computed for the complete analyses with $(q, n) = (q, n)_{\rm ML}$, and the dashed lines are those for the first iteration with $(q, n) = (1, 0)$. The shifts from the first to the third iteration steps are very small, but for all four component separation codes there is a consistent shift towards slightly higher amplitude for the third iteration, which also brings about an increase in the significance.

The amplitude of the dipole modulation changes with smoothing scale, as can be seen in figure 6.2b for maps from component separation with `Commander`, where the smoothing scales follow a rainbow coloring from red at 5° to purple at 10°. The amplitude becomes lower and less significant when increasing the smoothing scale from 5° to 8°, but increases again for smoothing scales 9° and 10°. Hanson and Lewis [92] pointed out that there is no obvious explanation for this trend, but it may be related to how cosmic variance affects the fluctuation for these smoothing scales. As before, the dashed lines correspond to the marginals for the analyses with $(q, n) = (1, 0)$ and the solid lines for that of the complete analyses. All six smoothing scales exhibit a shift towards higher amplitude and significance for the full analyses, that tend to be stronger for low $A_{\rm ML}$ results. The amplitude measurements

[3]The `Commander` map used had the galactic plane area that is inside the mask filled in by a constrained realization.

Data set	FWHM [°]	A	(l, b) [°]	q	n
Commander	5	$0.078^{+0.020}_{-0.021}$	$(227, -15) \pm 19$	0.95 ± 0.02	0.08 ± 0.04
NILC	5	$0.069^{+0.020}_{-0.021}$	$(226, -16) \pm 22$	0.95 ± 0.02	0.08 ± 0.04
SEVEM	5	$0.066^{+0.021}_{-0.021}$	$(227, -16) \pm 24$	0.95 ± 0.02	0.08 ± 0.04
SMICA	5	$0.065^{+0.021}_{-0.021}$	$(226, -17) \pm 24$	0.95 ± 0.02	0.08 ± 0.04
Commander	6	$0.076^{+0.024}_{-0.025}$	$(223, -16) \pm 25$	0.96 ± 0.03	0.10 ± 0.05
NILC	6	$0.062^{+0.025}_{-0.026}$	$(223, -19) \pm 38$	0.95 ± 0.03	0.09 ± 0.05
SEVEM	6	$0.060^{+0.025}_{-0.026}$	$(225, -19) \pm 40$	0.95 ± 0.03	0.10 ± 0.05
SMICA	6	$0.058^{+0.025}_{-0.027}$	$(223, -21) \pm 43$	0.96 ± 0.03	0.09 ± 0.05
Commander	7	$0.062^{+0.028}_{-0.030}$	$(223, -08) \pm 45$	0.96 ± 0.03	0.11 ± 0.06
NILC	7	$0.055^{+0.029}_{-0.030}$	$(225, -10) \pm 53$	0.96 ± 0.03	0.10 ± 0.06
SEVEM	7	$0.055^{+0.029}_{-0.030}$	$(226, -10) \pm 54$	0.96 ± 0.03	0.10 ± 0.06
SMICA	7	$0.048^{+0.029}_{-0.029}$	$(226, -11) \pm 58$	0.96 ± 0.03	0.08 ± 0.06
Commander	8	$0.043^{+0.032}_{-0.029}$	$(218, -15) \pm 62$	0.97 ± 0.03	0.11 ± 0.06
NILC	8	$0.049^{+0.032}_{-0.031}$	$(223, -16) \pm 59$	0.96 ± 0.03	0.08 ± 0.06
SEVEM	8	$0.050^{+0.032}_{-0.031}$	$(223, -15) \pm 60$	0.95 ± 0.03	0.09 ± 0.06
SMICA	8	$0.041^{+0.032}_{-0.029}$	$(225, -16) \pm 63$	0.95 ± 0.03	0.07 ± 0.06
Commander	9	$0.068^{+0.035}_{-0.037}$	$(210, -24) \pm 52$	0.96 ± 0.04	0.11 ± 0.07
NILC	9	$0.076^{+0.035}_{-0.037}$	$(216, -25) \pm 45$	0.95 ± 0.04	0.08 ± 0.07
SEVEM	9	$0.078^{+0.035}_{-0.037}$	$(215, -24) \pm 43$	0.95 ± 0.04	0.09 ± 0.07
SMICA	9	$0.070^{+0.035}_{-0.037}$	$(216, -25) \pm 50$	0.94 ± 0.04	0.07 ± 0.07
Commander	10	$0.092^{+0.037}_{-0.040}$	$(215, -29) \pm 38$	0.95 ± 0.04	0.10 ± 0.08
NILC	10	$0.098^{+0.037}_{-0.039}$	$(217, -29) \pm 33$	0.93 ± 0.04	0.07 ± 0.08
SEVEM	10	$0.103^{+0.037}_{-0.039}$	$(217, -28) \pm 30$	0.93 ± 0.04	0.08 ± 0.08
SMICA	10	$0.094^{+0.037}_{-0.040}$	$(218, -29) \pm 37$	0.93 ± 0.04	0.06 ± 0.07
*WMAP*5 ILC	4.5	0.072 ± 0.022	$(224, -22) \pm 24$	N/A	N/A
*WMAP*3 ILC	9	0.114	$(225, -27)$	N/A	N/A

Table 6.1: Best-fit dipole modulation parameters after three iterations, $(A, \hat{\boldsymbol{p}})_{\mathrm{ML},2}$, and the best-fit power spectrum parameters, $(q, n) = (q, n)_{\mathrm{ML}}$ for the six smoothing scales for each of the four component separation maps.

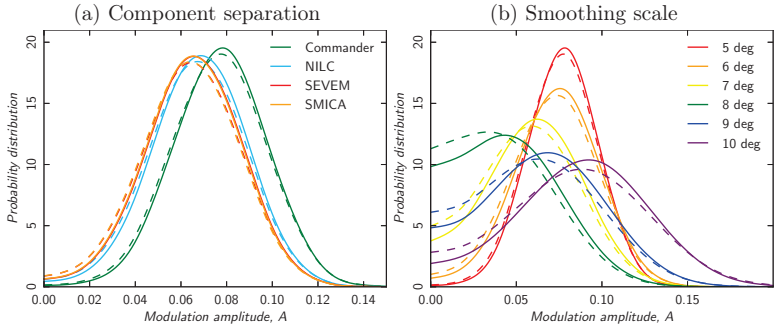


Figure 6.2: Normalized marginal probability distributions of the modulation amplitudes for a) the different component separation techniques at the 5° smoothing scale, and b) the six smoothing scales for the `Commander` component separation. The solid lines are the marginals for the full analyses with $(q, n) = (q, n)_{\rm ML}$ and the dashed are for the intermediate $(q, n) = (1, 0)$ analyses.

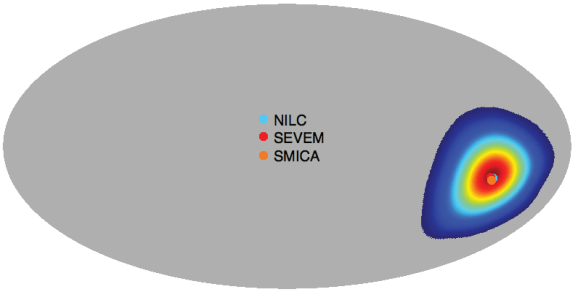
for all smoothing scales and component separation techniques appear to be consistent with $A \sim 0.07$.

The final likelihood distribution for the 5° `Commander` solution, smoothed with a 15° Gaussian beam and upgraded to $N_{\rm side} = 128$ is shown in figure 6.3a, where the likelihoods that are more than 2σ from the preferred direction are not shown. The best-fit directions obtained for the equivalent `NILC`, `SEVEM` and `SMICA` solutions are show with, respectively, blue, red and orange dots, on top of `Commander`'s distribution. There is clearly excellent agreement between these four directions and therefore for the large scale structures in the component separated maps.

The best-fit directions for the `Commander` maps with different smoothing scales are shown in figure 6.3b, where the underlying distribution is the same as before, namely that for `Commander` with 5° smoothing with a 2σ cutoff. The dots represent the best-fit directions for the other smoothing scales following the same rainbow inspired coloring scheme as before. As can be seen there is some drift between the preferred directions, but all points are well within the 2σ boundary, with the largest deviation at $\sim 1\sigma$ for the 9° smoothing.

The shift in the best-fit direction between the first iteration with $(q, n) = (1, 0)$ and third iteration with $(q, n) = (q, n)_{\rm ML}$ is to move slightly further from the galactic disc and to reduce the 2σ contour. In other words the $\hat{\boldsymbol{p}}_{\rm ML,2}$ direction is further from the galactic plane than the $\hat{\boldsymbol{p}}_{\rm ML}$ direction, and it's standard deviation decreases.

(a) Component separation



(b) Smoothing scale

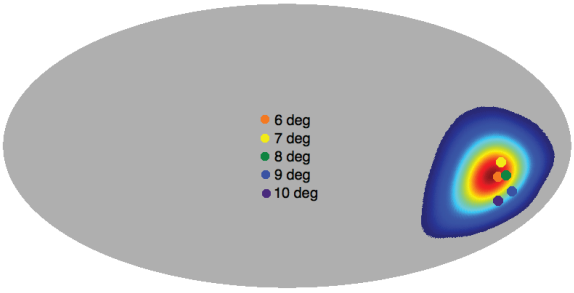


Figure 6.3: The best-fit modulation direction for a) the different component separation techniques at the 5° smoothing scale, and b) the six smoothing scales for the `Commander` component separation, where the underlying distribution is that for `Commander` smoothed to 5° with a cutoff at 2σ.

6.3.2 Power spectrum amplitude and tilt

The power spectrum amplitude and tilt parameters are strongly correlated, as can be seen in figure 6.4 for a smoothing scale of 5° for the different component separations in green, blue, red and orange for `Commander`, `NILC`, `SEVEM` and `SMICA`, respectively. The innermost contour captures the area close to the best-fit value and the outer three represent the 1σ, 2σ and 3σ levels. The black $\times$ is the parameter values corresponding to the fiducial model with $(q, n) = (1, 0)$. The best-fit values for these large scale analyses push the power spectrum towards a smaller amplitude and larger tilt, creating a tension with the fiducial model at a 1.7σ level in this case. This tension has been noted in the low-ℓ likelihood analysis by the *Planck* Collaboration [73], and was also seen in the *COBE*-DMR results [66, 67].

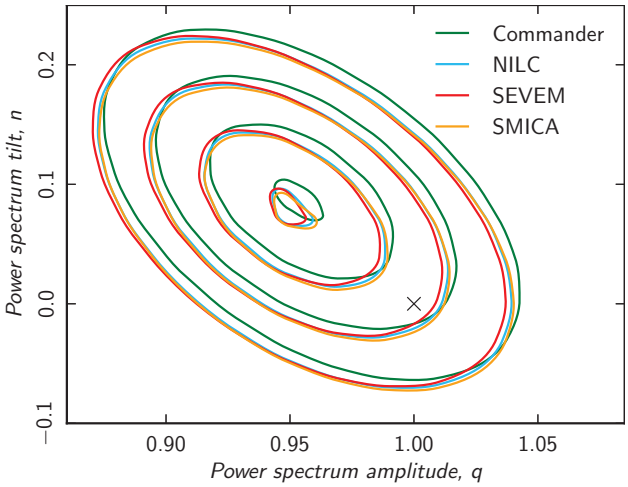


Figure 6.4: Contours for the power spectrum amplitude and tilt parameters for `Commander`, `NILC`, `SEVEM` and `SMICA` maps smoothed to 5° in, respectively, green, blue, red and orange, where the × corresponds to the fiducial power spectrum parameters $(q, n) = (1, 0)$. The innermost contour represents the area close to the best-fit values and the outer three contours are the 1σ, 2σ and 3σ contours.

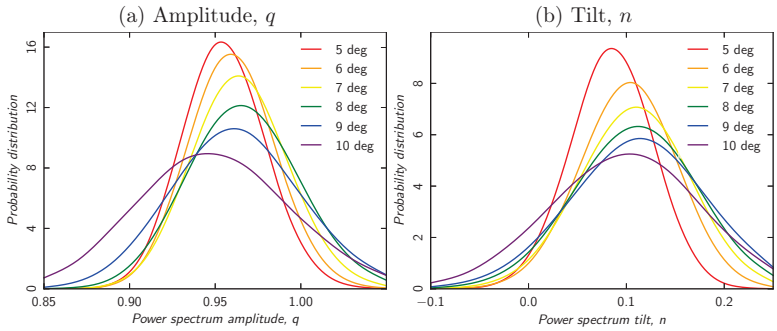


Figure 6.5: Normalized marginals probability distributions for the different smoothing scales for a) the power spectrum amplitude, q, and b) the power spectrum tilt, n, for the full analyses with $(q, n) = (q, n)_{\rm ML}$.

Figures 6.5a and 6.5b show the marginal distributions of the power spectrum amplitude and tilt, for the different smoothing scales of the Commander component separated map. The same trend of lower amplitude and higher tilt are visible here, but the significance of the deviation falls with increasing smoothing scale as a result of the broadening of the distribution. Furthermore, the values of the best-fit q and n increases from 5° to 8° and then decreases again. This trend is most likely related to that for the shift in best-fit modulation amplitude.

6.3.3 Significance

The 3 degrees of freedom log-likelihood differences for Commander, NILC, SEVEM and SMICA at each smoothing scale are shown in Figure 6.6, in green, blue, red and orange, respectively. The horizontal dashed lines show the 1 degree of freedom significance levels for 1σ, 2σ and 3σ, which correspond to a $\Delta \ln \mathcal{L}$ of $1.76, 4.01$ and 7.08, respectively. The values for the 3 dof log-likelihood difference and the corresponding 1 dof significance are shown in table 6.2 for each smoothing scale and component separation method.

Data set	**$\Delta \ln \mathcal{L}$**						**Significance** $[\sigma]$					
	5°	6°	7°	8°	9°	10°	5°	6°	7°	8°	9°	10°
Commander	8.8	6.4	4.0	2.1	3.3	4.5	3.5	2.8	2.0	1.2	1.7	2.2
NILC	7.1	4.7	3.4	2.5	3.9	5.0	3.0	2.3	1.7	1.4	1.9	2.3
SEVEM	6.7	4.6	3.3	2.5	4.0	5.4	2.9	2.2	1.7	1.4	2.0	2.5
SMICA	6.6	4.2	2.8	2.0	3.4	4.6	2.9	2.1	1.5	1.1	1.8	2.2

Table 6.2: Log-likelihood difference and significance of the six smoothing scales for each of the four component separation maps.

There is a clear overall trend for the codes with the significance falling steadily between smoothing scales of 5° and 8°, and then increasing again for 9° and 10°. This is expected since the equivalent trend is observed for the best-fit modulation amplitude, and power spectrum parameters. No satisfactory explanation has been found to explain this phenomenon, but with a maximum deviation of 0.7σ at low scales there is overall relatively good agreement between the different component separation techniques. The significance of Commander at the low smoothing scales sticks out from the rest, which might be linked to the usage of a constrained realization map for Commander.

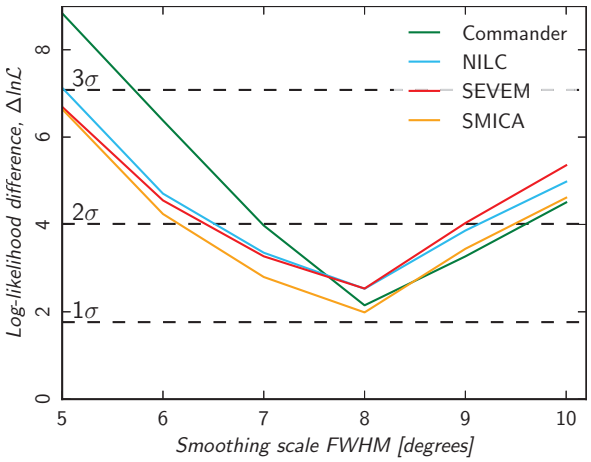


Figure 6.6: Plot of log-likelihood difference at each smoothing scale for `Commander`, `NILC`, `SEVEM` and `SMICA` in, respectively, green, blue, red and orange, where the dashed horizontal represent, from bottom to top, the 1σ, 2σ and 3σ levels.

6.4 Way forward

A number of improvements can be made to the existing algorithm for the dipole modulation model and the data being analyzed. In the current version of the algorithm the A, q and n grids are defined by the user, and the starting and ending points were chosen to fit the 5° smoothing scale case. This means the grid does not necessarily fit the other smoothing scales very well, as can be seen in figures 6.2b, 6.5a and 6.5b where the shift in amplitude and the broadening of the marginals with increasing smoothing scale pushes the marginal distributions past the grid boundaries. In order to capture the entire marginal distribution the starting and ending points should be adapted to each smoothing scale, which could be done in a variety of different ways. One way of doing this is by setting up a pre-computation stage for the parameters with a large grid containing only a few grid points, thereby covering the parameter space in few computations. The next stage uses the marginal distribution of the points from the first stage to set starting and ending points, for example by splining the marginal distribution to find the mean, μ, and the standard deviation, σ, and setting the grid edges at $\mu \pm 4\sigma$. Another, more complicated, way is to use adaptive grids for these parameters like for the dipole modulation direction. For this method the grid is also large to ensure the entire distribution is covered, and for the first stage there are only a few grid points. The difference arises in the continuation

of the computation, which in the case of the adaptive method is that grid points with likelihoods above a certain threshold are pushed to the next gridding level. Additional grid points are added half way between the grid points that survived, and only the new grid points go through likelihood evaluations. Once a sufficient resolution has been reached the gridding is complete, e.g. when there are at least 7 grid points covering $\mu \pm 4\sigma$, which ensures that the number of points covering this region is between 7 and 13.

The dipole modulation implemented here is a very simple scale-invariant phenomenological model, but the results from the analysis that extend to smaller scales (see section 5.5.1 of the paper) indicate that the amplitude falls off at higher multipoles. The modulation model should therefore be extended to more complicated classes of scale-dependent models, for example

$$A(\ell) = \alpha \left(\frac{\ell}{\ell_A} \right)^{\beta}, \tag{6.6}$$

where ℓ_A is the pivot multipole, and α and β are the amplitude of the modulation at ℓ_A and multipole dependence factor, respectively.

Another effect investigated by the *Planck* Collaboration, which was published in the March 2013 release was the impact of the Doppler effect caused by our motion with respect to the CMB radiation rest frame on the CMB map [93]. This was found to cause Doppler boosting in the direction of the dipole, which does not only affect the dipole but also higher order multipoles, and is accompanied by aberrations that distort the shapes and sizes of the temperature anisotropies in the direction of the dipole and the opposite direction. The dipole modulation analyses were performed on maps where the boosting effect was still present which probably pulled the resulting dipole modulation direction slightly towards the CMB dipole. Therefore, analyses of component separated maps that have been deboosted, i.e. where the Doppler effect have been removed, would improve the accuracy of the dipole modulation analysis.

Chapter 7

Paper IV: CMB Likelihood Approximation for Banded Probability Distributions

Together with the release of the data the *WMAP* team provided the likelihood evaluation code[1] used for parameter estimation. This code computes the likelihood of a model (represented by a set of parameters) by how well the proposal power spectrum this parameter set produces, C_ℓ, compares to the data, $\boldsymbol{d}$. In other words it computes the likelihood of a parameter set representing a model or the probability of observed data given this model, $\mathcal{L}(\theta|H) = P(d|\theta, H)$. The most common way of doing this is through maximum-likelihood evaluations over the parameter space of the model, and the CMB power spectrum likelihood is simply given by

$$\mathcal{L}(C_\ell) \equiv P(\boldsymbol{d}|C_\ell) \propto \frac{\mathrm{e}^{-\frac{1}{2}\boldsymbol{d}^T(\boldsymbol{S}(C_\ell)+\boldsymbol{N})^{-1}\boldsymbol{d}}}{\sqrt{|\boldsymbol{S}(C_\ell)+\boldsymbol{N}|}}, \tag{7.1}$$

when the signal and noise covariance matrices, $\boldsymbol{S}$ and $\boldsymbol{N}$, are Gaussians.

The *WMAP* likelihood evaluation utilizes a hybrid method to compute the full likelihood where a Gibbs sampling method based on a Blackwell-Rao estimator is used at large scale ($\ell_{\mathrm{min}} < \ell < \ell_{\mathrm{mid}}$) and at smaller scales ($\ell_{\mathrm{mid}} < \ell < \ell_{\mathrm{max}}$) the Master algorithm is used [94]. The low-ℓ Gibbs likelihood, $\mathcal{L}_G(L)$ and high-ℓ Master likelihood, $\mathcal{L}_M(H)$ likelihoods are merged at log-likelihood level to give the total likelihood as:

$$\ln \mathcal{L}_{\mathrm{tot}}(F) = \ln \mathcal{L}_G(L) + \ln \mathcal{L}_M(H) \qquad \text{with} \tag{7.2}$$

$$\ln \mathcal{L}_M(H) = \ln \mathcal{L}_M(H|\boldsymbol{C}_F), \tag{7.3}$$

[1] Available on LAMBDA (`http://lambda.gsfc.nasa.gov/`).

where the three ranges and the high-ℓ covariance matrix are given by

$$\begin{aligned} F &= \ell_{\rm min} < \ell < \ell_{\rm max}, && (7.4)\\ L &= \ell_{\rm min} < \ell < \ell_{\rm mid}, && (7.5)\\ H &= \ell_{\rm mid} < \ell < \ell_{\rm max} \qquad \text{and} && (7.6)\\ \boldsymbol{C}_F &= \boldsymbol{C}(\ell_{\rm min} < \ell < \ell_{\rm max}). && (7.7) \end{aligned}$$

This implementation covers correlations at all scales, including the off-diagonal terms between the high- and low-ℓ computations as can be seen in figure 7.1a, however there is also the option of a sharp transition which ignores these off-diagonal correlations between the high- and low-ℓ regions, as in figure 7.1b, which the *Planck* Collaboration has implemented [73].

7.1 Transition region

An alternative to these approaches is to introducing a transition region for the range $\ell_{\rm low}+1<\ell<\ell_{\rm high}-1$ where for simplicity $\ell_{\rm high}=\ell_{\rm mid}$. In this approach the low-ℓ computation ends at $\ell_{\rm high}-1$ and the high-ℓ computation starts at $\ell_{\rm low}+1$, creating an overlap region. Only the correlations between the $\ell<\ell_{\rm low}$ and $\ell>\ell_{\rm high}$ regions are ignored, and this omission is assumed to be correct provided the transition region is wide enough to make these correlations negligible. The total log-likelihood is then computed as

$$\ln\mathcal{L}_{\rm tot}(F)=\ln\mathcal{L}_G(L,T)+\ln\mathcal{L}_M(T,H|\boldsymbol{C}_{T,H})-\ln\mathcal{L}(T), \tag{7.8}$$

where the covariance matrix for the Master evaluations has been recomputed to only include the relevant correlations across T and H regions. The ranges for the different regions have been modified to

$$\begin{aligned} L &= \ell_{\rm min} < \ell < \ell_{\rm low}, && (7.9)\\ T &= \ell_{\rm low}+1 < \ell < \ell_{\rm high}-1 \qquad \text{and} && (7.10)\\ H &= \ell_{\rm high} < \ell < \ell_{\rm max}. && (7.11) \end{aligned}$$

This is pictorially represented by figure 7.1c where the likelihood for the transition region $\mathcal{L}(T)$ can be computed using either the Gibbs or Master approach, however only the Master approach was implement here, in which case $\ln\mathcal{L}(T)=\ln\mathcal{L}_M(T|\boldsymbol{C}_T)$, where the covariance matrix for the transition region, $\boldsymbol{C}_T$, is recomputed to contain the correlations for this region only.

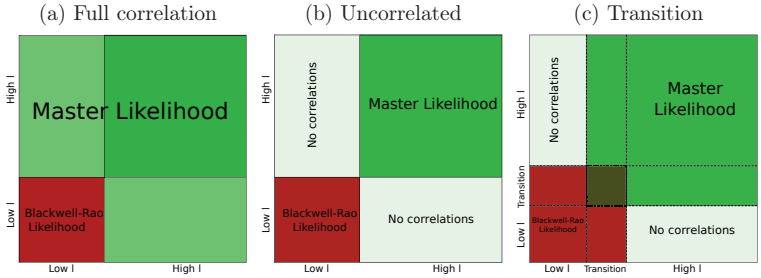


Figure 7.1: Region definitions for the combination of likelihoods with the low-ℓ using Blackwell-Rao estimator in red and high-ℓ Master computation in green, for a) *WMAP*-like combination with full correlation, b) *Planck*-like sharp transition with no off-diagonal correlation and c) a transition region with an overlap and no low- and high-ℓ corelation. Figures courtesy of E. Gjerløw.

7.2 Results

The cosmological parameter estimation code, CosmoMC, was run with the likelihood combinations defined by the three methods mentioned above, namely the fully correlated *WMAP*-like, the uncorrelated *Planck*-like and the transition region definitions. These runs were performed using the standard six parameter ΛCDM model with $\Omega_b h^2$, $\Omega_{DM} h^2$, θ, τ, n_s and $\log[10^{10} A_s]$ under the same computational conditions and with the same input parameters.

The resulting chains were analyzed using the accompanying software to produce the best-fit parameter values in table 7.1 and the marginal distributions shown in figure 7.2, where the approaches of a transition region, full correlation and uncorrelated high- and low-ℓ are shown in solid black, dashed red and dotted blue lines, respectively. Table 7.1 also shows the shift in best-fit parameters from the original *WMAP* code to the other two, in terms of the standard deviation obtained for the *WMAP*-like code, with a maximum shift of 0.06σ. Clearly, the difference between these three methods is negligible, with the conclusion that the addition of a transition region makes very little difference to the best-fit parameters as compared to the original *WMAP* likelihood code or to the *Planck*-like version with no off-diagonal correlations. This also means that the contribution from the off-diagonal terms around the low- and high-ℓ turn-over are so insignificant that their exclusion from the likelihood computation does not affect parameter estimation.

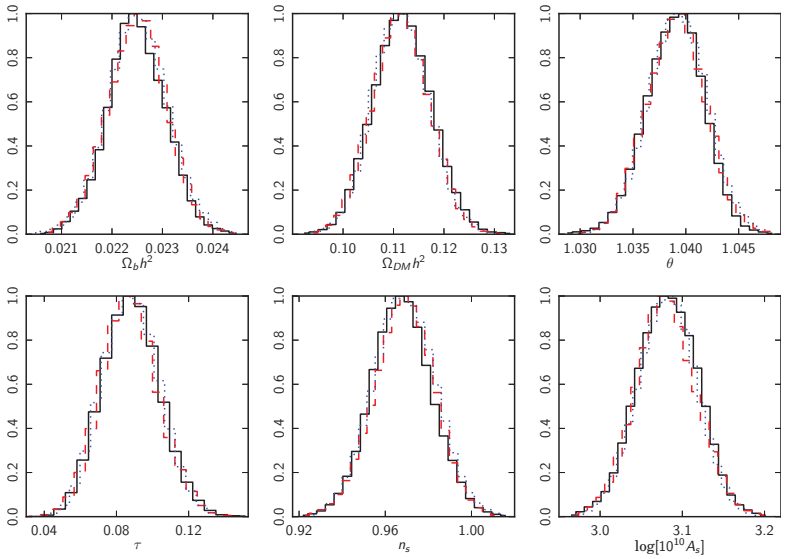


Figure 7.2: Normalized marginal distributions for the six parameter ΛCDM model for the transition approach (black solid line), the fully correlated *WMAP*-like combination (dashed red) and the *Planck*-like approach with no off-diagonal correlations at low- and high-ℓ (dotted blue).

Parameter	**Original WMAP**	**Sharp transition**		**Transition region**	
	Constraint	Constraint	Shift $[\sigma]$	Constraint	Shift $[\sigma]$
$\Omega_b h^2$	0.0225 ± 0.0006	0.0225 ± 0.0006	0.02	0.0225 ± 0.0006	0.02
$\Omega_m h^2$	0.111 ± 0.005	0.111 ± 0.005	0.01	0.112 ± 0.006	0.05
θ	1.039 ± 0.003	1.039 ± 0.003	0.04	1.039 ± 0.003	0.05
τ	0.088 ± 0.015	0.088 ± 0.015	0.04	0.088 ± 0.015	0.05
n_s	0.969 ± 0.013	0.969 ± 0.014	0.03	0.968 ± 0.014	0.06
$\log[10^{10} A_s]$	3.08 ± 0.04	3.08 ± 0.03	0.03	3.08 ± 0.04	0.05

Table 7.1: Best-fit parameter values for the original *WMAP* likelihood code, as compared to those obtained when implementing a sharp transition or a transition region, with the accompanying shifts in terms of the standard deviation of the original *WMAP* approach.

7.3 Way forward

The transition approach can be extended to more than 3 separate regions, where for n separate regions the total probability is given by

$$\begin{aligned} P(\{\boldsymbol{\theta}\}) &= P(\theta_1, \theta_2, ..., \theta_n) = \prod_{i=1}^{n-1} P(\theta_i|\theta_{i+1}, ..., \theta_n) \cdot P(\theta_n) && (7.12) \\ &\approx \prod_{i=1}^{n-1} P(\theta_i|\theta_{i+1}) \cdot P(\theta_n) && (7.13) \\ &= \prod_{i=1}^{n-1} \frac{P(\theta_i, \theta_{i+1})}{P(\theta_{i+1})} \cdot P(\theta_n) = \frac{\prod_{i=1}^{n-1} P(\theta_i, \theta_{i+1})}{\prod_{i=2}^{n-1} P(\theta_i)} && (7.14) \end{aligned}$$

where the second line is only valid if the correlations between regions θ_i and θ_{i+2} for $i = \{1 : n-2\}$ are negligible, which is called a tridiagonally banded probability. The third line comes from relating the joint probability to the conditional and prior probabilities through the relationship $P(x, y) = P(x|y) \cdot P(y)$. However, since the results for three regions did not differ noticeably from that of two fully correlated regions, the extension to more regions will probably not make any significant improvements.

Even though, the result mentioned in the previous section indicate that there is nothing gained by adding a transition region in the computation of the total likelihood, for completeness the effect of computing the likelihood in the transition region with the Gibbs based Blackwell Rao estimator should be checked, thus $\mathcal{L}(T) = \mathcal{L}_G(T)$.

Furthermore, for consistency the same analyses should be made with the 9-year data from the *WMAP* satellite, as well as checking how the addition of complete off-diagonal correlations or correlations through a transition region would affect the likelihood computations by the *Planck* Collaboration.

Chapter 8

Paper V: Grid-based Exploration of Cosmological Parameter Space with Snake

Cosmological models are described using a set of parameters, which can be used to compare and contrast different models, and to see how well a particular model fits the available data through likelihood evaluation, as introduced in chapter 7. These parameters usually represent some physical process or important quantity in the Universe and therefore enable us to compare a theoretical model to the actual Universe as observed by *WMAP*, *Planck*, the Sloan Digital Sky Survey (SDSS) [95] or other experiments. The aim of parameter estimation is to recover the set of parameters that fit the experimental data best.

The most commonly utilized code in the cosmological community for this purpose is CosmoMC [96, 97], which investigates parameter space through Markov-Chain Monte Carlo (MCMC) sampling, and by default uses a Metropolis-Hasting approach. These sampling methods perform random walks to sample multi-dimensional parameter space, where the position of the current parameter point and a proposal density function are used to select the next point. The probability of keeping a new point is directly related to its likelihood value, thus this sampling method produces long chains of randomly positioned samples with a higher density of points in regions where the underlying likelihood distribution is greater.

Alternatively, the parameter space can be sampled with a grid-based sampling method, which has a very simple implementation since the possible sets of parameters are limited to fixed points on a pre-defined grid, based on the start position and size of grid-cells. Furthermore, having data on such a regularized grid makes extraction of conditional distributions and computation of Bayesian evidence for model comparrison trivial, and allows for thorough investigation of the tails of the underlying distribution.

The major disadvantage of a grid-based sampling technique is the "curse of dimensionality". For each additional dimension the number of samples required goes up by a factor x, the number of grid points for accurate sampling in one dimension, thus for n dimensions x^n number of samples are required. The typical number of samples in cosmological analyses is of the order of 10^5, which corresponds to the number of samples for a five parameter model if $x = 10$ or if the number of grid points per dimension is decreased to 6 or 7 a grid-based method should easily sample a six parameter model.

8.1 Snake algorithm

The candidate developed and tested a grid-based sampling algorithm in order to investigate how well a relatively simple grid-based method would compare to the MCMC sampling of CosmoMC. Snake differs from an ordinary grid-based method due to the implementation of a threshold cutoff for evaluations, where grid points with likelihood values below this cutoff will not be investigated. Moreover, Snake is written to fit into the CosmoMC package and uses the existing routines within CosmoMC to compute the likelihood for a given set of parameters, thus only the sampling method has been swapped out. Due to the difference between the samples MCMC and grid-based sampling methods produce the outputs and post-processing routines for Snake also differ from those of CosmoMC.

Before going through the algorithm itself let us define the terminology for Snake:

The grid. The user selects the parameters of the start position, $\boldsymbol{\theta}_0$, and the size of a grid-cell, $\boldsymbol{\Delta\theta}$, thereby setting up the structure of the virtual grid. For Snake the grid is infinite in size although only part of it will be investigated.

Grid points For easy book-keeping each grid point is assigned an integer, i, representing that this point was the ith point to be investigated. The position of any point on the grid is referred to through an integer vector, $\boldsymbol{k}_i$, that describes its position with respect to the starting position in terms of grid-cell sizes, thus $\boldsymbol{\theta}_i = \boldsymbol{\theta}_0 + \boldsymbol{k}_i \cdot \boldsymbol{\Delta\theta}$.

The surface. Grid points fall into three possible categories: external, surface or internal points. External points are those that have not yet been investigated, of which there are an infinite number; surface points are grid points that have been investigated but that have at least one unexplored neighbor; and internal points are those for which the point itself and all its neighbors have been investigated. The three categories comprise of points that are outside, on or inside the surface, respectively.

The overall and surface maximum likelihoods. The point on the grid with the overall highest likelihood is said to be the peak point of the likelihood distribution and is represented by the integer $i_{\rm peak}$. The surface point with the highest likelihood is represented by the integer $i_{\rm S}$, which continually changes as the surface evolves.

The repository. Visited parameter points are stored in a data structure called a repository, where each row in a two-dimensional dynamic list defines a point on the grid. In addition to the integer i each data row contains the point's $\boldsymbol{k}_i$ vector, likelihood value $\mathcal{L}_i$, the integers corresponding to its neighbors, and a logical flag specifying whether the point is on the surface.

Dictionaries. In order to have rapid book-keeping standard C++ map template dictionaries have been implemented, which store the combination of two values; the key and the mapped value. The advantage of these maps is that the mapped values are easily and quickly accessed from the key. Two dictionaries are implemented in Snake, one for keeping track of which iteration a visited grid point corresponds to, $\boldsymbol{k}_i \to i$; and another for the iteration index of a particular likelihood value, $\mathcal{L}_i \to i$. The latter map is sorted according to descending likelihood ensuring that the point with the highest likelihood is that at the top of the list. Once a surface point becomes internal its dictionary entries are removed so that only surface points exist in the dictionaries[1].

Threshold cutoff. The most basic grid-based method sets up a fixed grid and computes the likelihood at each grid-point, which would be all the points inside a square/rectangle for two-dimensional parameter space. Defining a cutoff-threshold[2] ensures that Snake only investigates neighbors of grid points that have likelihoods above a certain limit, where the excluded points do not contribute any useful information. The threshold thereby restricts the investigated two-dimensional parameter space to a circle/ellipse. A higher threshold allows the tails to be investigated further, but also increases the computational cost, and therefore a balance has to be found between a sufficiently high threshold and a reasonable computational cost.

[1] The $i_{\rm S}$ point is therefore always the entry at the top of the likelihood map.

[2] Convergence is reached when the log-likelihood difference between the surface and overall likelihoods, given by $\log \mathcal{L}(i_{\rm S}) - \log \mathcal{L}(i_{\rm peak})$, exceeds the threshold.

The Snake algorithm can be summarized as follows:

1. *Initialization.* Set up the virtual grid, the repository structure and initialize the C++ dictionaries. Compute the likelihood of the starting point, and insert the required information into the repository and dictionaries.

2. *Neighbor.* Consider the surface point with the highest likelihood value, located at the top of the $\mathcal{L}_i \to i$ map, which has index i_{S}, and randomly pick one of its unexplored neighbors. Evaluate the likelihood of this new point.

3. *Surface update.* Use the parameter dictionary, $\boldsymbol{k}_i \to i$, to find the indices corresponding to the neighbors of the new point. If any of these neighbors or the new point no longer have any unexplored neighbors, the corresponding surface flag(s) is/are set to false.

4. *Update dictionaries:* If the new point is classed as a surface point the relevant information is inserted into the two dictionaries. Any of the neighbors that have become internal points are removed from the dictionaries.

5. *Update repository.* Insert the new point into the repository structure and update the relevant information for its neighbor points.

6. *Convergence check.* Snake has converged if the maximum number of samples has been reached or the likelihoods of all surface points are below the threshold, and the investigation of parameter space will subsequently be terminated. Otherwise Snake loops back to step 2 and performs another iteration.

The most time consuming part of an iteration for a particular parameter point is the evaluation of its likelihood, and running several evaluations at the same time would therefore be beneficial. Snake is not locked to Markov chains, which have to be a certain length before they are deemed to have converged, but can simultaneously perform as many evaluations as there are available processors. However, in the initial phase the number of unexplored neighbors is limited and therefore the speed-up achieved with increased number of processors is only nearly perfect. A master–slave parallelization strategy has been implemented, where one single master is the administrator that deals with the "paperwork" and assigns work to a number of slaves. More specifically, the slaves do the hard task of evaluating the likelihoods for the parameter sets provided by the master and return the results to the master upon completion. The master keeps the repository and dictionaries up to date, provides external grid points to idle slaves and gathers the likelihood results.

However, this parallelization scheme adds an algorithmic complication that was not included in the algorithm summary above, namely how the master goes about sending multiple parameter points for evaluation, how it knows when the answers are ready and how to deal with the parameter points that have been sent for evaluation but for which the slave has not yet returned a likelihood value. In the simple summary the new point is chosen and the likelihood evaluation is performed, after which the surface, dictionaries and repository are updated and then a new point is chosen. However, when dealing with multiple slaves it is crucial that the master and slaves don't waste time by waiting for each other, and the parallelization has to be optimized. This is acheived by following a specific procedure:

- *Send work.* The master sends parameter sets to idle slaves for likelihood evaluation as long as i) available jobs exist (i.e. $i_{\rm S}$ point has unexplored neighbors), ii) there are slaves without an assigned task, and iii) the threshold has not been reached.

- *Neighbor checks.* Immediately after a parameter set is assigned to a slave the point is entered into the parameter dictionary ($\boldsymbol{k}_i \rightarrow i$), with the purpose of enabling neighbor checks in order to update the neighbor status in the repository and dictionaries, and also to update the $i_{\rm S}$ point. This therefore enables the master to keep track of which neighbors to the $i_{\rm S}$ point that can be assigned to idle slaves.

- *Receive reply.* The master now settles down to wait for the first slave to return with an evaluated likelihood, and subsequently checks the surface status and enters the required information in the repository and dictionaries.

- *Prepare next round.* Before going back to assigning work the master updates the overall and surface maximum likelihood points, $i_{\rm peak}$ and $i_{\rm S}$, but exits the iterations if the threshold or maximum number of samples has been reached.

The algorithm has been thoroughly tested to ensure that it yields correct results for multi-dimensional likelihood distributions and for distributions with several local maxima. Furthermore, an analysis of the accuracy of results with increasing number of dimensions was performed for an uncorrelated Gaussian multi-dimensional likelihood distribution. The outcome of the investigation was that for a maximum number of samples of 10^6, Snake is capable of investigating parameter space with 12 dimensions to high accuracy.

8.2 Results for *WMAP* 7-year data

The Snake method of investigating parameter space was applied to the 7-year *WMAP* data in order to estimate the best-fit parameter values for the concordance ΛCDM model, with the 6 standard parameters $\Omega_b h^2$, $\Omega_{DM} h^2$, θ, τ, n_s and $\log[10^{10} A_s]$ [98]. The threshold cutoff was set to -6.0 and the code was run on a total of 50 processors, thus 49 slaves were crunching away at likelihood evaluations simultaneously. An equivalent run was performed with the standard Metropolis-Hastings MCMC sampling in CosmoMC with an MPI convergence criterion of 0.03, and also using 50 processors thereby allowing 50 chains to be run.

These two runs resulted in wall times of 1.24 hours and 1.42 hours for Snake and CosmoMC, respectively, and post-analyses produced the best-fit parameters and standard deviations shown in table 8.1. The column titled 'shift' quantizes how many σs the best-fit parameters differ by, and with the greatest discrepancy of 0.08σ the agreement between the two methods is excellent. Figure 8.1 shows the normalized marginal distributions for the six ΛCDM parameters, where Snake corresponds to the dotted red lines and CosmoMC to the solid black lines. The level of agreement here is also clear as the two histograms are more or less indistinguishable.

Parameter	**CosmoMC**	**Snake**	**Shift [σ]**
$\Omega_b h^2$	$0.02252^{+0.00055}_{-0.00056}$	$0.02252^{+0.00057}_{-0.00056}$	0
$\Omega_{DM} h^2$	$0.1110^{+0.0055}_{-0.0054}$	$0.1107^{+0.0055}_{-0.0054}$	0.06
θ	1.039 ± 0.003	1.039 ± 0.003	0
τ	$0.08849^{+0.00632}_{-0.00754}$	$0.08758^{+0.01558}_{-0.01426}$	0.08
n_s	$0.9682^{+0.0138}_{-0.0136}$	$0.9681^{+0.0139}_{-0.0138}$	0.07
$\log[10^{10} A_s]$	$3.082^{+0.034}_{-0.035}$	3.080 ± 0.035	0.06

Table 8.1: Numerical results for two equivalent runs with sampling by the standard Metropolis-Hastings setup and by Snake in the CosmoMC framework for the standard six-parameter ΛCDM model, together with the difference in best-fit parameters in terms of the standard deviations.

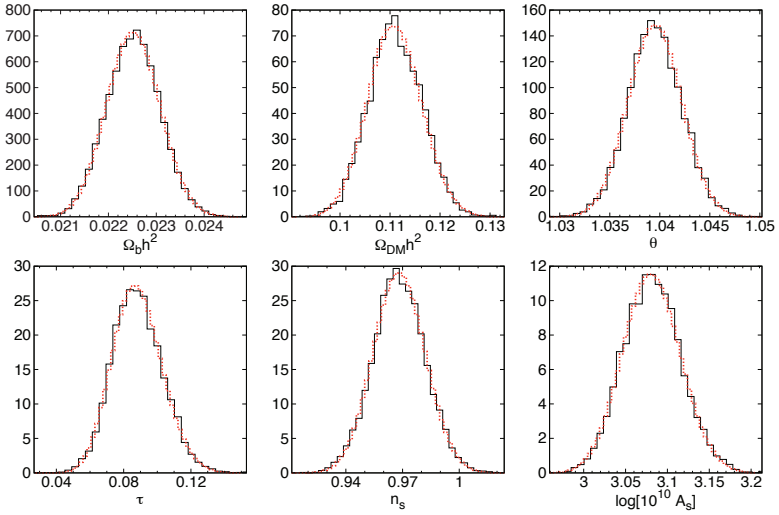


Figure 8.1: Normalized marginal distributions of the six standard ΛCDM model parameters for equivalent runs with Snake and the standard CosmoMC sampling in dotted red and solid black lines, respectively.

The panels in figure 8.2 show four examples of normalized two-dimensional likelihood distributions for pairs of parameters in the ΛCDM model, where the other four parameters have been integrated out. The color scheme represents the likelihood distribution where areas in deepest blue (value 0) are those outside the threshold cutoff and the darkest red (value 1) corresponds to the maximum likelihood. The skewness of a two-dimensional distribution displays the amount of correlation between the two parameters, giving an indication of how strongly pairs of parameters are affected by each other. A pair of parameters that are completely uncorrelated give a spherical/elliptical distribution that is perfectly aligned along the x and y axes. The distributions in the panels at the top are nearly perfectly spherical and are excellent examples of parameter pairs with negligible correlation. Strongly correlated parameter pairs result in distributions that are highly elongated and incline diagonally such that the elongation axis lies at an angle to the x and y angles. The bottom panels show distributions for strongly correlated parameter pairs, where the strength of the correlation increases with increasing flattening or decreasing compression[3], and an inclination angle of 45° maximizes the dependence.

[3]For an ellipse with the semi-major and semi-minor axes represented by a and b, the flattening is computed as $f = (a - b)/a$ and compression is given by b/a.

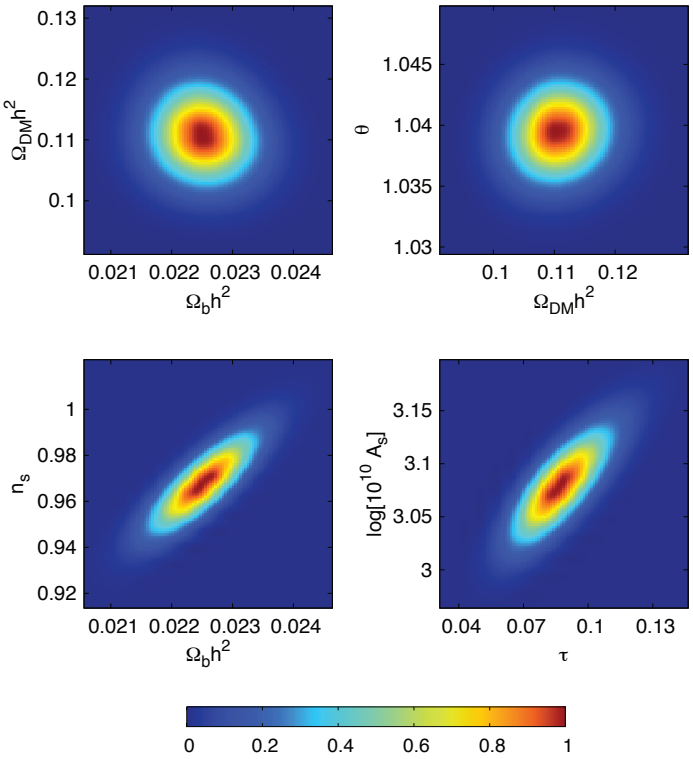


Figure 8.2: Normalized two-dimensional marginal distributions for pairs of parameters for the six parameter ΛCDM model where the darkest red, with value 1, corresponds to the maximum likelihood and the darkest blue to the area outside the threshold, with value 0. The skewness of the distribution indicates how correlated the parameters are, where a weak correlation gives horizontal/vertical ellipses/spheres like the top panels, and a strong correlation gives a flattened diagonal distribution, as in the lower panels.

8.3 Bayesian evidence

Comparrison of different models can be performed through Bayesian evidence, which is the probability of data $\boldsymbol{d}$ given model H with parameters $\boldsymbol{\theta}$. The Bayesian evidence over parameter space Ω is given by

$$P(\boldsymbol{d}|H) = \int_{\Omega} P(\boldsymbol{d}, \boldsymbol{\theta}|H) d\boldsymbol{\theta} = \int_{\Omega} P(\boldsymbol{d}|\boldsymbol{\theta}, H) P(\boldsymbol{\theta}|H) d\boldsymbol{\theta}, \tag{8.1}$$

where $P(\boldsymbol{d}, \boldsymbol{\theta}|H)$ is the joint probability of data $\boldsymbol{d}$ and parameters $\boldsymbol{\theta}$ for a given model H, and $P(\boldsymbol{d}|\boldsymbol{\theta}, H) = \mathcal{L}(\boldsymbol{\theta}|H)$ is the likelihood of parameter set $\boldsymbol{\theta}$. $P(\boldsymbol{\theta}|H)$ is the prior distribution, which is assumed to be uniform and given by $P(\theta_i|H) = 1/L_i$ for each parameter, where L_i is the range of significant likelihood contributions for parameter i that has been set to $\mu_i \pm 4\sigma_i$. The combined priors for N_{par}-dimensional space is are given by

$$P(\boldsymbol{\theta}|H) = \prod_{i=1}^{N_{\text{par}}} \frac{1}{L_i} = \frac{1}{V_{\text{range}}}, \tag{8.2}$$

where V_{range} is the N_{par}-dimensional volume enclosed by the L_i ranges. The Bayesian evidence takes into consideration the model's complexity, such that a model with an extra parameter that fits the data better does not necessarily give a better evidence, because the added complexity is penalized.

Computing Bayesian evidence for a grid-based sampling regime is more straight forward than for an MCMC method like CosmoMC, which has samples distributed irregularly in parameter space. For a grid-based sampling regime the integration of likelihood over the entire parameter space in equation (8.1) can be simplified to a summation of likelihoods multiplied by a fixed cell volume, $V_{\text{cell}} = \prod_{i=1}^{N(par)} \Delta\theta_i$. The Bayesian evidence for N samples with Snake is given by

$$P(\boldsymbol{d}|H) = \frac{V_{\text{cell}}}{V_{\text{range}}} \sum_{i=1}^{N} \mathcal{L}_i, \tag{8.3}$$

requiring just a simple summation of likelihood values for the samples and the trivial computations of cell and range volumes.

The evidence of a model by itself is not very informative with regards to how well this model describes the data. However, two models, represented by H_1 and H_2, can be compared through their evidences in order to decide which model fits better. The standard way of performing this model comparison is through the difference in log-evidences, given by

$$\delta \log E = \log E_1 - \log E_2, \tag{8.4}$$

and by setting boundaries using the Jeffrey's scale [99, 100, 101], which is given by

$$\delta \log E > \begin{cases} 1 & \text{evidence for } E_1 \text{ is substantial} \\ 2.5 & \text{evidence for } E_1 \text{ is strong} \\ 5 & \text{evidence for } E_1 \text{ is decisive.} \end{cases}$$

This scale is by no means set in stone and should be used more as a guideline for model comparison rather than as a fixed recipe (see [101] for a recent review of this).

As an illustration the Bayesian evidence is used to check if the standard six parameter ΛCDM model described above is a significantly better fit to the *WMAP* 7-year data than the simpler five parameter model with a non-varying scale factor ($n_s = 1$). The two models are represented by H_1 and H_2, respectively. The L_i ranges for the five common parameters ($\Omega_b h^2$, $\Omega_{DM} h^2$, θ, τ, and $\log[10^{10} A_s]$) are fixed to $\mu \pm 4\sigma$ as computed from analysis of the five-parameter model H_2, and that for the additional n_s parameter in the six parameter model is computed as $\mu \pm 4\sigma$ from the results of the H_1 model. The resulting evidences are $E_1 = -3743.21$ and $E_2 = -3744.62$ giving a substantial log-evidence difference of 1.41 in favor of the standard six parameter ΛCDM model. This is in agreement with previous findings [102], and indicates that adding the n_s parameter makes a significantly better fit to the data, even when taking into consideration the increased complexity.

8.4 Way forward

Even though Snake utilizes the CosmoMC framework it is currently not perfectly merged into CosmoMC with regards to data format of the output and the accompanying analysis software, which would obviously make Snake more desirable. The difference in the way the two methods sample parameter space needs to be taken into account by making a post-processing software that can deal with both MCMC samples and regularized grid samples, and producing the same type of numerical results and plots.

The Snake sampling method relies on preselecting the grid-cell sizes for each parameter, which suffers from a "Goldilocks syndrome". A step size that is too small will result in a high density of grid points, which is computationally expensive and the maximum number of allowed steps will probably cut the analysis short before completion. On the other hand, a too large step size means that Snake does not sample the underlying likelihood distribution thoroughly enough to be useful for parameter estimation purposes. Therefore, the step sizes for all parameter have to be "just right", which requires a lot of fine tuning. It would be prudent to extend the method to an adaptive grid-method, similar to that used for the dipole modulation direction (section 6.2) and described as improvements for the other parameters. The

method would be the same as discussed previously, where the initial grid is coarse with only a few grid points covering a large parameter range, which is subdivided until a sufficiently high resolution is achieved. Only grid points with likelihoods above a cutoff value would survive to the next grid resolution, but in this case only grid points half way between the grid points of the previous level are sent for likelihood evaluation. The effect of this method would be to zoom in on the maximum likelihood area of the distribution, while keeping the number of evaluations to a minimum.

Additionally, one could take a page out of the CosmoMC handbook and implement n_c different "chains", which would mean having n_c masters that each have their repository and dictionary structures, and a set of slaves to do their bidding. The chains should be initialized at different places in parameter space to get around the problem of multiple maxima not being found if the trough likelihood is below the threshold. However, care has to be taken to ensure that the chains use the same infinitely sized virtual grid structure, and furthermore, if two chains meet the two repositories, dictionaries and slaves should be merged.

Chapter 9

Bibliography

[1] H. S. Leavitt. 1777 variables in the Magellanic Clouds. *Annals of Harvard College Observatory*, 60:87–108, 1908.

[2] E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929.

[3] H. Bondi and T. Gold. The Steady-State Theory of the Expanding Universe. *Monthly Notices of the Royal Astronomical Society*, 108:252, 1948.

[4] F. Hoyle. A New Model for the Expanding Universe. *Monthly Notices of the Royal Astronomical Society*, 108:372, 1948.

[5] A. Friedmann. Über die Krümmung des Raumes. *Zeitschrift fur Physik*, 10:377–386, 1922.

[6] A. Friedmann. Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes. *Zeitschrift fur Physik*, 21:326–332, December 1924.

[7] G. Lemaître. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Societe Scietifique de Bruxelles*, 47:49–59, 1927.

[8] G. Gamow. Expanding Universe and the Origin of Elements. *Physical Review*, 70:572–573, October 1946.

[9] R. A. Alpher, H. Bethe, and G. Gamow. The Origin of Chemical Elements. *Physical Review*, 73:803–804, April 1948.

[10] M. Ryle and R. W. Clarke. An examination of the steady-state model in the light of some recent observations of radio sources. *Monthly Notices of the Royal Astronomical Society*, 122:349, 1961.

[11] R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic Black-Body Radiation. *Astrophysical Journal*, 142:414–419, July 1965.

[12] A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal*, 142:419–421, July 1965.

[13] S. Dodelson. *Modern cosmology*. 2003.

[14] J. A. Peacock. *Cosmological Physics*. January 1999.

[15] B. Gumjudpai. Introductory Overview of Modern Cosmology. *ArXiv Astrophysics e-prints*, May 2003, arXiv:astro-ph/0305063.

[16] D. D. Reid, D. W. Kittell, E. E. Arsznov, and G. B. Thompson. The picture of our universe: A view from modern cosmology. *ArXiv Astrophysics e-prints*, September 2002, arXiv:astro-ph/0209504.

[17] R. H. Brandenberger. Modern Cosmology and Structure Formation. In J. F. Donoghue, editor, *CP Violation and the Limits of the Standard Model*, page 551, 1995, arXiv:astro-ph/9411049.

[18] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. *ArXiv e-prints*, December 2012, 1212.5226.

[19] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XVI. Cosmological parameters. *ArXiv e-prints*, March 2013, 1303.5076.

[20] A. Einstein. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847.*, pages 844–847, 1915.

[21] A. Einstein. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 142-152.*, pages 142–152, 1917.

[22] F. W. Dyson, A. S. Eddington, and C. Davidson. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Royal Society of London Philosophical Transactions Series A*, 220:291–333, 1920.

[23] E. F. Taylor and J. A. Wheeler. *Exploring black holes : introduction to general relativity*. 2000.

[24] H. P. Robertson. Kinematics and World-Structure. *Astrophysical Journal*, 82:284, November 1935.

[25] A. G. Walker. On the formal comparison of Milne's kinematical system with the systems of general relativity. *Monthly Notices of the Royal Astronomical Society*, 95:263–269, January 1935.

[26] S. Perlmutter, G. Aldering, M. della Valle, S. Deustua, R. S. Ellis, S. Fabbro, A. Fruchter, G. Goldhaber, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, R. A. Knop, C. Lidman, R. G. McMahon, P. Nugent, R. Pain, N. Panagia, C. R. Pennypacker, P. Ruiz-Lapuente, B. Schaefer, and N. Walton. Discovery of a supernova explosion at half the age of the universe. *Nature*, 391:51, January 1998, arXiv:astro-ph/9712212.

[27] B. P. Schmidt, N. B. Suntzeff, M. M. Phillips, R. A. Schommer, A. Clocchiatti, R. P. Kirshner, P. Garnavich, P. Challis, B. Leibundgut, J. Spyromilio, A. G. Riess, A. V. Filippenko, M. Hamuy, R. C. Smith, C. Hogan, C. Stubbs, A. Diercks, D. Reiss, R. Gilliland, J. Tonry, J. Maza, A. Dressler, J. Walsh, and R. Ciardullo. The High-Z Supernova Search: Measuring Cosmic Deceleration and Global Curvature of the Universe Using Type IA Supernovae. *Astrophysical Journal*, 507: 46–63, November 1998, arXiv:astro-ph/9805200.

[28] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, and J. Tonry. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astronomical Journal*, 116: 1009–1038, September 1998, arXiv:astro-ph/9805201.

[29] F. Hoyle and R. J. Tayler. The Mystery of the Cosmic Helium Abundance. *Nature*, 203:1108–1110, September 1964.

[30] P. J. E. Peebles. Primordial Helium Abundance and the Primordial Fireball. II. *Astrophysical Journal*, 146:542, November 1966.

[31] R. V. Wagoner, W. A. Fowler, and F. Hoyle. On the Synthesis of Elements at Very High Temperatures. *Astrophysical Journal*, 148:3, April 1967.

[32] D. J. Fixsen, G. Hinshaw, C. L. Bennett, and J. C. Mather. The Spectrum of the Cosmic Microwave Background Anisotropy from the Combined COBE FIRAS and DMR Observations. *Astrophysical Journal*, 486:623, September 1997, arXiv:astro-ph/9704176.

[33] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. *ArXiv e-prints*, December 2012, 1212.5225.

[34] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.

[35] F. Zwicky. On the Masses of Nebulae and of Clusters of Nebulae. *Astrophysical Journal*, 86:217, October 1937.

[36] V. C. Rubin, W. K. J. Ford, and N. . Thonnard. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *Astrophysical Journal*, 238:471–487, June 1980.

[37] D. Walsh, R. F. Carswell, and R. J. Weymann. 0957 + 561 A, B - Twin quasistellar objects or gravitational lens. *Nature*, 279:381–384, May 1979.

[38] A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Physical Review D*, 23:347–356, January 1981.

[39] K. Sato. First-order phase transition of a vacuum and the expansion of the Universe. *Monthly Notices of the Royal Astronomical Society*, 195:467–479, May 1981.

[40] A. D. Linde. A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. *Physics Letters B*, 108:389–393, February 1982.

[41] A. Albrecht and P. J. Steinhardt. Cosmology for grand unified theories with radiatively induced symmetry breaking. *Physical Review Letters*, 48:1220–1223, April 1982.

[42] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XXII. Constraints on inflation. *ArXiv e-prints*, March 2013, 1303.5082.

[43] A. H. Guth and S.-Y. Pi. Fluctuations in the new inflationary universe. *Physical Review Letters*, 49:1110–1113, October 1982.

[44] S. W. Hawking. The development of irregularities in a single bubble inflationary universe. *Physics Letters B*, 115:295–297, September 1982.

[45] A. A. Starobinsky. Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. *Physics Letters B*, 117:175–178, November 1982.

[46] R. K. Sachs and A. M. Wolfe. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *Astrophysical Journal*, 147:73, January 1967.

[47] Y. B. Zeldovich and R. A. Sunyaev. The Interaction of Matter and Radiation in a Hot-Model Universe. *Astrophysics & Space Science*, 4: 301–316, July 1969.

[48] R. A. Sunyaev and Y. B. Zeldovich. The interaction of matter and radiation in the hot model of the Universe, II. *Astrophysics & Space Science*, 7:20–30, April 1970.

[49] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XIX. The integrated Sachs-Wolfe effect. *ArXiv e-prints*, March 2013, 1303.5079.

[50] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XX. Cosmology from Sunyaev-Zeldovich cluster counts. *ArXiv e-prints*, March 2013, 1303.5080.

[51] N. W. Boggess, J. C. Mather, R. Weiss, C. L. Bennett, E. S. Cheng, E. Dwek, S. Gulkis, M. G. Hauser, M. A. Janssen, T. Kelsall, S. S. Meyer, S. H. Moseley, T. L. Murdock, R. A. Shafer, R. F. Silverberg, G. F. Smoot, D. T. Wilkinson, and E. L. Wright. The COBE mission - Its design and performance two years after launch. *Astrophysical Journal*, 397:420–429, October 1992.

[52] J. C. Mather, D. J. Fixsen, R. A. Shafer, C. Mosier, and D. T. Wilkinson. Calibrator Design for the COBE Far-Infrared Absolute Spectrophotometer (FIRAS). *Astrophysical Journal*, 512:511–520, February 1999, arXiv:astro-ph/9810373.

[53] D. J. Fixsen. The Temperature of the Cosmic Microwave Background. *Astrophysical Journal*, 707:916–920, December 2009, 0911.1955.

[54] D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, and E. L. Wright. The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set. *Astrophysical Journal*, 473:576, December 1996, arXiv:astro-ph/9605054.

[55] S. Padin, M. C. Shepherd, J. K. Cartwright, R. G. Keeney, B. S. Mason, T. J. Pearson, A. C. S. Readhead, W. A. Schaal, J. Sievers, P. S. Udomprasert, J. K. Yamasaki, W. L. Holzapfel, J. E. Carlstrom, M. Joy, S. T. Myers, and A. Otarola. The Cosmic Background Imager. *NaturePublications of the Astronomical Society of the Pacific*, 114:83–97, January 2002, arXiv:astro-ph/0110124.

[56] A. C. Taylor, M. E. Jones, J. R. Allison, E. Angelakis, J. R. Bond, L. Bronfman, R. Bustos, R. J. Davis, C. Dickinson, J. Leech, B. S. Mason, S. T. Myers, T. J. Pearson, A. C. S. Readhead, R. Reeves, M. C. Shepherd, and J. L. Sievers. The Cosmic Background Imager 2. *Monthly Notices of the Royal Astronomical Society*, 418:2720–2729, December 2011, 1108.3950.

[57] E. M. Leitch, C. Pryke, N. W. Halverson, J. Kovac, G. Davidson, S. LaRoque, E. Schartman, J. Yamasaki, J. E. Carlstrom, W. L. Holzapfel, M. Dragovan, J. K. Cartwright, B. S. Mason, S. Padin, T. J. Pearson, A. C. S. Readhead, and M. C. Shepherd. Experiment Design and First Season Observations with the Degree Angular Scale Interferometer. *Astrophysical Journal*, 568:28–37, March 2002, arXiv:astro-ph/0104488.

[58] M. C. Runyan, P. A. R. Ade, R. S. Bhatia, J. J. Bock, M. D. Daub, J. H. Goldstein, C. V. Haynes, W. L. Holzapfel, C. L. Kuo, A. E. Lange, J. Leong, M. Lueker, M. Newcomb, J. B. Peterson, C. Reichardt, J. Ruhl, G. Sirbi, E. Torbet, C. Tucker, A. D. Turner, and D. Woolsey. ACBAR: The Arcminute Cosmology Bolometer Array Receiver. *Astrophysical Journal, Supplement*, 149:265–287, December 2003, arXiv:astro-ph/0303515.

[59] A. Kosowsky. The Atacama Cosmology Telescope. *New Astronomy Review*, 47:939–943, December 2003, arXiv:astro-ph/0402234.

[60] J. Ruhl, P. A. R. Ade, J. E. Carlstrom, H.-M. Cho, T. Crawford, M. Dobbs, C. H. Greer, N. w. Halverson, W. L. Holzapfel, T. M. Lanting, A. T. Lee, E. M. Leitch, J. Leong, W. Lu, M. Lueker, J. Mehl, S. S. Meyer, J. J. Mohr, S. Padin, T. Plagge, C. Pryke, M. C. Runyan, D. Schwan, M. K. Sharp, H. Spieler, Z. Staniszewski, and A. A. Stark. The South Pole Telescope. In C. M. Bradford, P. A. R. Ade, J. E. Aguirre, J. J. Bock, M. Dragovan, L. Duband, L. Earle, J. Glenn, H. Matsuhara, B. J. Naylor, H. T. Nguyen, M. Yun, and J. Zmuidzinas, editors, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5498 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 11–29, October 2004, arXiv:astro-ph/0411122.

[61] P. de Bernardis, P. A. R. Ade, R. Artusa, J. J. Bock, A. Boscaleri, B. P. Crill, G. De Troia, P. C. Farese, M. Giacometti, V. V. Hristov, A. Iacoangeli, A. E. Lange, A. T. Lee, S. Masi, L. Martinis, P. V. Mason, P. D. Mauskopf, F. Melchiorri, L. Miglio, T. Montroy, C. B. Netterfield, E. Pascale, F. Piacentini, P. L. Richards, J. E. Ruhl, and F. Scaramuzzi. Mapping the CMB sky: THE BOOMERanG experiment. *New Astronomy Review*, 43:289–296, July 1999, arXiv:astro-ph/9911461.

[62] A. T. Lee, P. Ade, A. Balbi, J. Bock, J. Borrill, A. Boscaleri, B. P. Crill, P. de Bernardis, H. Del Castillo, P. Ferreira, K. Ganga, S. Hanany, V. Hristov, A. H. Jaffe, A. E. Lange, P. Mauskopf, C. B. Netterfield, S. Oh, E. Pascale, B. Rabii, P. L. Richards, J. Ruhl, G. F. Smoot, and C. D. Winant. MAXIMA: an experiment to measure temperature anisotropy in the cosmic microwave background. In L. Maiani, F. Melchiorri, and N. Vittorio, editors, *3K cosmology*, volume 476 of *American Institute of Physics Conference Series*, pages 224–236, May 1999, arXiv:astro-ph/9903249.

[63] B. Rabii, C. D. Winant, J. S. Collins, A. T. Lee, P. L. Richards, M. E. Abroe, S. Hanany, B. R. Johnson, P. Ade, A. Balbi, J. J. Bock, J. Borrill, R. Stompor, A. Boscaleri, E. Pascale, P. de Bernardis, P. G. Ferreira, V. V. Hristov, A. E. Lange, A. H. Jaffe, C. B. Netterfield, G. F. Smoot, and J. H. P. Wu. MAXIMA: A balloon-borne cosmic microwave background anisotropy experiment. *Review of Scientific Instruments*, 77(7):071101, July 2006, arXiv:astro-ph/0309414.

[64] A. Benoît, P. Ade, A. Amblard, R. Ansari, E. Aubourg, J. Bartlett, J.-P. Bernard, R. S. Bhatia, A. Blanchard, J. J. Bock, A. Boscaleri, F. R. Bouchet, A. Bourrachot, P. Camus, F. Couchot, P. de Bernardis, J. Delabrouille, F.-X. Désert, O. Doré, M. Douspis, L. Dumoulin, X. Dupac, P. Filliatre, K. Ganga, F. Gannaway, B. Gautier, M. Giard, Y. Giraud-Héraud, R. Gispert, L. Guglielmi, J.-C. Hamilton, S. Hanany, S. Henrot-Versillé, V. V. Hristov, J. Kaplan, G. Lagache, J.-M. Lamarre, A. E. Lange, K. Madet, B. Maffei, D. Marrone, S. Masi, J. A. Murphy, F. Naraghi, F. Nati, G. Perrin, M. Piat, J.-L. Puget, D. Santos, R. V. Sudiwala, J.-C. Vanel, D. Vibert, E. Wakui, and D. Yvon. Archeops: a high resolution, large sky coverage balloon experiment for mapping cosmic microwave background anisotropies. *Astroparticle Physics*, 17:101–124, May 2002, arXiv:astro-ph/0106152.

[65] J. C. Mather, E. S. Cheng, D. A. Cottingham, R. E. Eplee, Jr., D. J. Fixsen, T. Hewagama, R. B. Isaacman, K. A. Jensen, S. S. Meyer, P. D. Noerdlinger, S. M. Read, L. P. Rosen, R. A. Shafer, E. L. Wright, C. L. Bennett, N. W. Boggess, M. G. Hauser, T. Kelsall, S. H. Moseley, Jr., R. F. Silverberg, G. F. Smoot, R. Weiss, and D. T. Wilkinson. Mea-

surement of the cosmic microwave background spectrum by the COBE FIRAS instrument. *Astrophysical Journal*, 420:439–444, January 1994.

[66] G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE differential microwave radiometer first-year maps. *Astrophysical Journal, Letters*, 396:L1–L5, September 1992.

[67] C. L. Bennett, A. J. Banday, K. M. Gorski, G. Hinshaw, P. Jackson, P. Keegstra, A. Kogut, G. F. Smoot, D. T. Wilkinson, and E. L. Wright. Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results. *Astrophysical Journal, Letters*, 464:L1, June 1996, arXiv:astro-ph/9601067.

[68] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. Healpix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *Astrophysical Journal*, 622(2):759, 2005, astro-ph/0409513v1. URL `http://stacks.iop.org/0004-637X/622/i=2/a=759`.

[69] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. VI. High Frequency Instrument data processing. *ArXiv e-prints*, March 2013, 1303.5067.

[70] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XIV. Zodiacal emission. *ArXiv e-prints*, March 2013, 1303.5074.

[71] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XXIII. Isotropy and Statistics of the CMB. *ArXiv e-prints*, March 2013, 1303.5083.

[72] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XII. Component separation. *ArXiv e-prints*, March 2013, 1303.5072.

[73] Planck collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XV. CMB power spectra and likelihood. *ArXiv e-prints*, March 2013, 1303.5075.

[74] H. K. Eriksen, A. J. Banday, K. M. Górski, and P. B. Lilje. On foreground removal from the wilkinson microwave anisotropy probe data by an internal linear combination method: Limitations and implications. *Astrophysical Journal*, 612:633–646, September 2004, arXiv:astro-ph/0403098.

[75] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission. *Astrophysical Journal, Supplement*, 148:97–117, September 2003, arXiv:astro-ph/0302208.

[76] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack. The WMAP First Year Source Catalog (WMAP1) (Bennett+, 2003). *VizieR Online Data Catalog*, 214:80097, September 2003.

[77] P. Vielva, E. Martínez-González, R. B. Barreiro, J. L. Sanz, and L. Cayón. Detection of Non-Gaussianity in the Wilkinson Microwave Anisotropy Probe First-Year Data Using Spherical Wavelets. *Astrophysical Journal*, 609:22–34, July 2004, arXiv:astro-ph/0310273.

[78] M. Cruz, E. Martínez-González, P. Vielva, and L. Cayón. Detection of a non-Gaussian spot in WMAP. *Monthly Notices of the Royal Astronomical Society*, 356:29–40, January 2005, arXiv:astro-ph/0405341.

[79] M. Tegmark, A. de Oliveira-Costa, and A. J. Hamilton. High resolution foreground cleaned CMB map from WMAP. *Physical Review D*, 68 (12):123523, December 2003, arXiv:astro-ph/0302496.

[80] A. de Oliveira-Costa, M. Tegmark, M. Zaldarriaga, and A. Hamilton. Significance of the largest scale CMB fluctuations in WMAP. *Physical Review D*, 69(6):063516, March 2004, arXiv:astro-ph/0307282.

[81] C.-G. Park. Non-Gaussian signatures in the temperature fluctuation observed by the Wilkinson Microwave Anisotropy Probe. *Monthly Notices of the Royal Astronomical Society*, 349:313–320, March 2004, arXiv:astro-ph/0307469.

[82] H. K. Eriksen, F. K. Hansen, A. J. Banday, K. M. Górski, and P. B. Lilje. Asymmetries in the Cosmic Microwave Background Anisotropy Field. *Astrophysical Journal*, 605:14–20, April 2004, arXiv:astro-ph/0307507.

[83] F. K. Hansen, P. Cabella, D. Marinucci, and N. Vittorio. Asymmetries in the Local Curvature of the Wilkinson Microwave Anisotropy Probe Data. *Astrophysical Journal, Letters*, 607:L67–L70, June 2004, arXiv:astro-ph/0402396.

[84] M. Cruz, L. Cayón, E. Martínez-González, P. Vielva, and J. Jin. The Non-Gaussian Cold Spot in the 3 Year Wilkinson Microwave Anisotropy Probe Data. *Astrophysical Journal*, 655:11–20, January 2007, arXiv:astro-ph/0603859.

[85] C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman. On the large-angle anomalies of the microwave sky. *Monthly Notices of the Royal Astronomical Society*, 367:79–102, March 2006, arXiv:astro-ph/0508047.

[86] C. L. Bennett, R. S. Hill, G. Hinshaw, D. Larson, K. M. Smith, J. Dunkley, B. Gold, M. Halpern, N. Jarosik, A. Kogut, E. Komatsu, M. Limon, S. S. Meyer, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Are There Cosmic Microwave Background Anomalies? *Astrophysical Journal, Supplement*, 192:17, February 2011, 1001.4758.

[87] M. Axelsson, Y. Fantaye, F. K. Hansen, A. J. Banday, H. K. Eriksen, and K. M. Gorski. Directional Dependence of ΛCDM Cosmological Parameters. *Astrophysical Journal, Letters*, 773:L3, August 2013, 1303.5371.

[88] C. Gordon, W. Hu, D. Huterer, and T. Crawford. Spontaneousisotropy breaking: A mechanism for CMB multipole alignments. *Physical Review D*, 72(10):103002, November 2005, arXiv:astro-ph/0509301.

[89] H. K. Eriksen, A. J. Banday, K. M. Górski, F. K. Hansen, and P. B. Lilje. Hemispherical Power Asymmetry in the Third-Year Wilkinson Microwave Anisotropy Probe Sky Maps. *Astrophysical Journal, Letters*, 660:L81–L84, May 2007, arXiv:astro-ph/0701089.

[90] J. Hoftuft, H. K. Eriksen, A. J. Banday, K. M. Górski, F. K. Hansen, and P. B. Lilje. Increasing Evidence for Hemispherical Power Asymmetry in the Five-Year WMAP Data. *Astrophysical Journal*, 699:985–989, July 2009, 0903.1229.

[91] F. K. Hansen, A. J. Banday, K. M. Górski, H. K. Eriksen, and P. B. Lilje. Power Asymmetry in Cosmic Microwave Background Fluctuations from Full Sky to Sub-Degree Scales: Is the Universe Isotropic? *Astrophysical Journal*, 704:1448–1458, October 2009, 0812.3795.

[92] D. Hanson and A. Lewis. Estimators for CMB statistical anisotropy. *Physical Review D*, 80(6):063004, September 2009, 0908.0963.

[93] Planck Collaboration, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2013 results. XXVII. Doppler boosting of the CMB: Eppur si muove. *ArXiv e-prints*, March 2013, 1303.5087.

[94] Eric Hivon, Krzysztof M. Górski, C. Barth Netterfield, Brendan P. Crill, Simon Prunet, and Frode Hansen. Master of the cosmic microwave background anisotropy power spectrum: A fast method for statistical analysis of large and complex cosmic microwave background data sets. *Astrophysical Journal*, 567(1):2, 2002, astro-ph/0105302v1. URL `http://stacks.iop.org/0004-637X/567/i=1/a=2`.

[95] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee,

J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda, and SDSS Collaboration. The Sloan Digital Sky Survey: Technical Summary. *Astronomical Journal*, 120: 1579–1587, September 2000, arXiv:astro-ph/0006396.

[96] Antony Lewis and Sarah Bridle. Cosmological parameters from cmb and other data: A monte carlo approach. *Physical Review D*, 66: 103511, Nov 2002, astro-ph/0205436v3. URL `http://link.aps.org/doi/10.1103/PhysRevD.66.103511`.

[97] Antony Lewis and Sarah Bridle. CosmoMC Notes. `http://cosmologist.info/notes/CosmoMC.pdF`.

[98] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolta, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophysical Journal, Supplement*, 192:18, February 2011, 1001.4538.

[99] A. Liddle, P. Mukherjee, and D. Parkinson. Model selection in cosmology. *Astronomy and Geophysics*, 47(4):040000–4, August 2006.

[100] R. Trotta. Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemporary Physics*, 49:71–104, March 2008, 0803.4089.

[101] S. Nesseris and J. García-Bellido. Is the Jeffreys' scale a reliable tool for Bayesian model comparison in cosmology? *Journal of Cosmology and Astroparticle Physics*, 8:36, August 2013, 1210.7652.

[102] D. Parkinson and A. R. Liddle. Application of Bayesian model averaging to measurements of the primordial power spectrum. *Physical Review D*, 82(10):103533, November 2010, 1009.1394.

Part III

Papers

Paper I

***Planck* 2013 results. XXIX. *Planck* catalogue of Sunyaev–Zeldovich sources**

Planck Collaboration

Submitted to Astronomy & Astrophysics

Astronomy & Astrophysics manuscript no. PSZcatalogue
March 22, 2013

Planck 2013 results. XXIX. *Planck* catalogue of Sunyaev–Zeldovich sources

Planck Collaboration: P. A. R. Ade[96], N. Aghanim[66*], C. Armitage-Caplan[101], M. Arnaud[79], M. Ashdown[76,7], F. Atrio-Barandela[21], J. Aumont[66], H. Aussel[79], C. Baccigalupi[94], A. J. Banday[107,11], R. B. Barreiro[73], R. Barrena[72], M. Bartelmann[105,85], J. G. Bartlett[1,74], E. Battaner[110], K. Benabed[67,104], A. Benoît[64], A. Benoit-Lévy[29,67,104], J.-P. Bernard[11], M. Bersanelli[40,56], P. Bielewicz[107,11,94], I. Bikmaev[24,3], J. Bobin[79], J. J. Bock[74,12], H. Böhringer[86], A. Bonaldi[75], J. R. Bond[10], J. Borrill[16,98], F. R. Bouchet[67,104], M. Bridges[76,7,70], M. Bucher[1], R. Burenin[97,89], C. Burigana[55,38], R. C. Butler[55], J.-F. Cardoso[80,1,67], P. Carvalho[7], A. Catalano[81,78], A. Challinor[70,76,13], A. Chamballu[79,18,66], R.-R. Chary[63], X. Chen[63], L.-Y Chiang[69], H. C. Chiang[32,8], G. Chon[86], P. R. Christensen[90,43], E. Churazov[85,97], S. Church[100], D. L. Clements[62], S. Colombi[67,104], L. P. L. Colombo[28,74], B. Comis[81], F. Couchot[77], A. Coulais[78], B. P. Crill[74,91], A. Curto[7,73], F. Cuttaia[55], A. Da Silva[14], H. Dahle[71], L. Danese[94], R. D. Davies[75], R. J. Davis[75], P. de Bernardis[39], A. de Rosa[55], G. de Zotti[52,94], J. Delabrouille[1], J.-M. Delouis[67,104], J. Démoclès[79], F.-X. Désert[59], C. Dickinson[75], J. M. Diego[73], K. Dolag[109,85], H. Dole[66,65], S. Donzelli[56], O. Doré[74,12], M. Douspis[66], X. Dupac[46], G. Efstathiou[70], T. A. Enßlin[85], H. K. Eriksen[71], F. Feroz[7], F. Finelli[55,57], I. Flores-Cacho[11,107], O. Forni[107,11], M. Frailis[54], E. Franceschi[55], S. Fromenteau[1,66], S. Galeotta[54], K. Ganga[1], R. T. Génova-Santos[72], M. Giard[107,11], G. Giardino[47], M. Gilfanov[85,97], Y. Giraud-Héraud[1], J. González-Nuevo[73,94], K. M. Górski[74,112], K. J. B. Grainge[7,76], S. Gratton[76,70], A. Gregorio[41,54], N, E. Groeneboom[71], A. Gruppuso[55], F. K. Hansen[71], D. Hanson[87,74,10], D. Harrison[70,76], A. Hempel[72,44], S. Henrot-Versillé[77], C. Hernández-Monteagudo[15,85], D. Herranz[73], S. R. Hildebrandt[12], E. Hivon[67,104], M. Hobson[7], W. A. Holmes[74], A. Hornstrup[19], W. Hovest[85], K. M. Huffenberger[111], G. Hurier[66,81], N. Hurley-Walker[7], T. R. Jaffe[107,11], A. H. Jaffe[62], W. C. Jones[32], M. Juvela[31], E. Keihänen[31], R. Keskitalo[26,16], I. Khamitov[102,24], T. S. Kisner[83], R. Kneissl[45,9], J. Knoche[85], L. Knox[34], M. Kunz[20,66,4], H. Kurki-Suonio[31,50], G. Lagache[66], A. Lähteenmäki[2,50], J.-M. Lamarre[78], A. Lasenby[7,76], R. J. Laureijs[47], C. R. Lawrence[74], J. P. Leahy[75], R. Leonardi[46], J. León-Tavares[48,2], J. Lesgourgues[103,93], C. Li[84,85], A. Liddle[95,30], M. Liguori[37], P. B. Lilje[71], M. Linden-Vørnle[19], M. López-Caniego[73], P. M. Lubin[35], J. F. Macías-Pérez[81], C. J. MacTavish[76], B. Maffei[75], D. Maino[40,56], N. Mandolesi[55,6,38], M. Maris[54], D. J. Marshall[79], P. G. Martin[10], E. Martínez-González[73], S. Masi[39], S. Matarrese[37], F. Matthai[85], P. Mazzotta[42], S. Mei[49,106,12], P. R. Meinhold[35], A. Melchiorri[39,58], J.-B. Melin[18], L. Mendes[46], A. Mennella[40,56], M. Migliaccio[70,76], S. Mitra[61,74], M.-A. Miville-Deschênes[66,10], A. Moneti[67], L. Montier[107,11], G. Morgante[55], D. Mortlock[62], D. Munshi[96], P. Naselsky[90,43], F. Nati[39], P. Natoli[38,5,55], N. P. H. Nesvadba[66], C. B. Netterfield[23], H. U. Nørgaard-Nielsen[19], F. Noviello[75], D. Novikov[62], I. Novikov[90], I. J. O'Dwyer[74], M. Olamaie[7], S. Osborne[100], C. A. Oxborrow[19], F. Paci[94], L. Pagano[39,58], F. Pajot[66], D. Paoletti[55,57], F. Pasian[54], G. Patanchon[1], T. J. Pearson[12,63], O. Perdereau[77], L. Perotto[81], Y. C. Perrott[7], F. Perrotta[94], F. Piacentini[39], M. Piat[1], E. Pierpaoli[28], D. Pietrobon[74], S. Plaszczynski[77], E. Pointecouteau[107,11], G. Polenta[5,53], N. Ponthieu[66,59], L. Popa[68], T. Poutanen[50,31,2], G. W. Pratt[79], G. Prézeau[12,74], S. Prunet[67,104], J.-L. Puget[66], J. P. Rachen[25,85], W. T. Reach[108], R. Rebolo[72,17,44], M. Reinecke[85], M. Remazeilles[66,1], C. Renault[81], S. Ricciardi[55], T. Riller[85], I. Ristorcelli[107,11], G. Rocha[74,12], C. Rosset[1], G. Roudier[1,78,74], M. Rowan-Robinson[62], J. A. Rubiño-Martín[72,44], C. Rumsey[7], B. Rusholme[63], M. Sandri[55], D. Santos[81], R. D. E. Saunders[7,76], G. Savini[92], D. Scott[27], M. D. Seiffert[74,12], E. P. S. Shellard[13], T. W. Shimwell[7], L. D. Spencer[96], J.-L. Starck[79], V. Stolyarov[7,76,99], R. Stompor[1], R. Sudiwala[96], R. Sunyaev[85,97], F. Sureau[79], D. Sutton[70,76], A.-S. Suur-Uski[31,50], J.-F. Sygnet[67], J. A. Tauber[47], D. Tavagnacco[54,41], L. Terenzi[55], L. Toffolatti[22,73], M. Tomasi[56], M. Tristram[77], M. Tucci[20,77], J. Tuovinen[88], M. Türler[60], G. Umana[51], L. Valenziano[55], J. Valiviita[50,31,71], B. Van Tent[82], L. Vibert[66], P. Vielva[73], F. Villa[55], N. Vittorio[42], L. A. Wade[74], B. D. Wandelt[67,104,36], M. White[33], S. D. M. White[85], D. Yvon[18], A. Zacchei[54], and A. Zonca[35]

(Affiliations can be found after the references)

Preprint online version: March 22, 2013

ABSTRACT

We describe the all-sky *Planck* catalogue of clusters and cluster candidates derived from Sunyaev–Zeldovich (SZ) effect detections using the first 15.5 months of *Planck* satellite observations. The catalogue contains 1227 entries, making it over six times the size of the *Planck* Early SZ (ESZ) sample and the largest SZ-selected catalogue to date. It contains 861 confirmed clusters, of which 178 have been confirmed as clusters, mostly through follow-up observations, and a further 683 are previously-known clusters. The remaining 366 have the status of cluster candidates, and we divide them into three classes according to the quality of evidence that they are likely to be true clusters. The *Planck* SZ catalogue is the deepest all-sky cluster catalogue, with redshifts up to about one, and spans the broadest cluster mass range from $(0.1 \text{ to } 1.6) \times 10^{15}$ $M_{\odot}$. Confirmation of cluster candidates through comparison with existing surveys or cluster catalogues is extensively described, as is the statistical characterization of the catalogue in terms of completeness and statistical reliability. The outputs of the validation process are provided as additional information. This gives, in particular, an ensemble of 813 cluster redshifts, and for all these *Planck* clusters we also include a mass estimated from a newly-proposed SZ-mass proxy. A refined measure of the SZ Compton parameter for the clusters with X-ray counter-parts is provided, as is an X-ray flux for all the *Planck* clusters not previously detected in X-ray surveys.

Key words. large-scale structure of Universe – Galaxies: clusters: general – Catalogs

* Corresponding author: N. Aghanim
nabila.aghanim@ias.u-psud.fr

1. Introduction

This paper, one of a set associated with the 2013 release of data from the *Planck*[1] mission (Planck Collaboration I 2013), describes the construction and properties of the *Planck* catalogue of SZ sources (PSZ).

Clusters of galaxies play several important roles in astrophysics and cosmology. As rare objects, their number density is especially sensitive to properties of the cosmological model such as the amplitude of primordial density perturbations (Peebles 1980), and their development with redshift probes the growth of cosmic structure, hence perhaps helping to distinguish between dark energy and modified gravity explanations for cosmic acceleration (e.g., see reviews by Borgani & Kravtsov 2009; Allen et al. 2011). The galaxies, hot gas and dark matter held in their gravitational potential wells provide a sample of the universal abundance of these components (e.g., Voit 2005), while the thermal state of the gas probes both the cluster formation mechanism and physical processes within the cluster such as cooling and energy-injection feedback (e.g., reviews by Fabian 2012; McNamara & Nulsen 2012). The study of the constituent galaxies, including the brightest cluster galaxies normally found at their centres, allows sensitive tests of galaxy formation models.

Because of these uses, there is considerable interest in developing large galaxy cluster catalogues that can be used for population and cosmological studies (e.g., Schuecker et al. 2003; Böhringer et al. 2004). Clusters are genuinely multi-wavelength objects that can be selected in several ways: optical/infrared (IR) imaging of the galaxy populations; X-ray imaging of bremsstrahlung radiation from the hot cluster gas; and through the Sunyaev–Zeldovich (SZ) effect (Sunyaev & Zeldovich 1972; Sunyaev & Zeldovich 1980) whereby scattering of cosmic microwave background (CMB) photons from that hot gas distorts the spectral shape of the CMB along lines of sight through clusters and groups.

Construction of cluster catalogues in the optical/IR and in the X-ray are relatively mature activities. The first large optical cluster survey is now over 50 years old (Abell 1958; Abell et al. 1989), and current catalogues constructed from the Sloan Digital Sky Survey data contain over a hundred thousand clusters (e.g., Koester et al. 2007; Wen et al. 2012). In X-rays, large samples first became available via *ROSAT* satellite observations (e.g., Vikhlinin et al. 1998; Böhringer et al. 2000; Gioia et al. 2003; Böhringer et al. 2004; Burenin et al. 2007; Ebeling et al. 2007), but also more recently for instance from dedicated or serendipitous survey with *XMM-Newton* (Pacaud et al. 2007; Fassbender et al. 2011; Takey et al. 2011; Mehrtens et al. 2012). Currently several thousand X-ray selected clusters are known (see for instance the meta-catalogue MCXC by Piffaretti et al. 2011). By contrast, although proposed about fifteen years ago (e.g., Barbosa et al. 1996; Aghanim et al. 1997), it is only very recently that SZ-selected samples have reached a significant size, with publication of samples containing several hundred clusters from the Early SZ (ESZ) catalogue from the *Planck* Satellite (Planck Collaboration VIII 2011), the South Pole Telescope (SPT, Reichardt et al. 2013) and the Atacama Cosmology Telescope (ACT, Hasselfield et al. 2013).

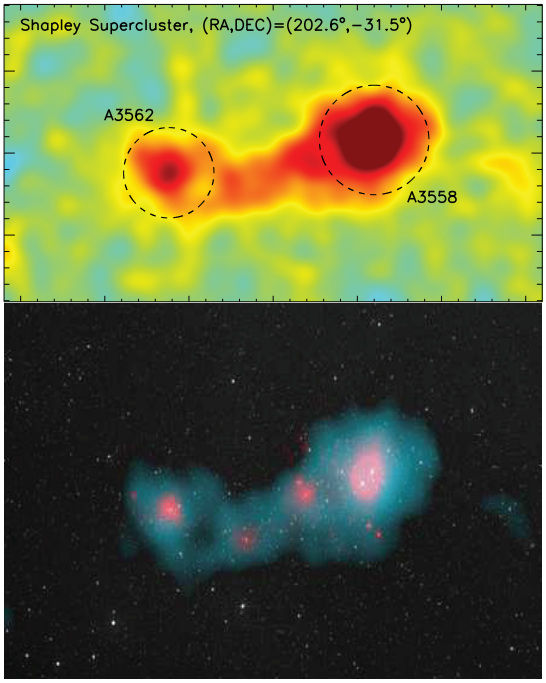


Fig. 1: The Shapley super-cluster as seen in the *Planck* survey. *Upper panel:* reconstructed thermal SZ map 3.2×1.8 square degrees from MILCA (Hurier et al. 2010). The dotted circles represent apertures of θ_{500} from the MCXC meta-catalogue around the resolved clusters. *Lower panel:* composite view of the optical from DSS images, X-rays from *ROSAT* (light pink) survey and from the thermal SZ effect as seen in *Planck* (cyan).

The usefulness of the different selection methods, particularly for cosmology, depends not just on the total number of clusters identified but also on how readily the selection function of the survey can be modelled, and on how well the observed cluster properties can be related to quantities such as the total cluster mass that are most readily predicted from theory (e.g., see Voit 2005). It has proven difficult to capitalize on the large size of optical/IR cluster samples because the observable, the number of galaxies in each cluster, exhibits large scatter with respect to the total cluster mass (e.g., Johnston et al. 2007). In this regard the X-ray selected samples are considerably more powerful, due to the tighter correlations of X-ray properties with mass (Arnaud et al. 2005; Vikhlinin et al. 2006; Pratt et al. 2009; Reichert et al. 2011; Maughan et al. 2012). Simulations predict that SZ-selected surveys may do even better, with a very tight relation between SZ signal and mass (e.g., da Silva et al. 2004; Motl et al. 2005; Nagai 2006; Wik et al. 2008; Aghanim et al. 2009; Angulo et al. 2012). Moreover, this relation, except at low redshifts, corresponds to a nearly redshift-independent mass limit, thus allowing such surveys to reach to high redshift and provide a strong lever arm on growth of structure.

We report on the construction and properties of the PSZ catalogue, which is to date the largest SZ-selected cluster catalogue and has value added through compilation of ancillary informa-

[1] *Planck* (http://www.esa.int/Planck) is a project of the European Space Agency (ESA) with instruments provided by two scientific consortia funded by ESA member states (in particular the lead countries France and Italy), with contributions from NASA (USA) and telescope reflectors provided by a collaboration between ESA and a scientific consortium led and funded by Denmark.

tion. It contains 1227 entries including many multiple systems, e.g., the Shapley super-cluster displayed in Fig. 1 together with a composite image. Of these 861 are confirmed, amongst which 178 are new discoveries, whilst amongst the 366 candidate clusters 54 are of high reliability (CLASS1 in our terminology), 170 are reliable, and the remaining 142 are in the lowest reliability class. In Sect. 2 we start with a description of the *Planck* data used to provide cluster candidates, and the two different methodologies (one of which has two independent implementations) used to carry out the extraction of the SZ sources. In Sect. 3 we provide a characterization of the PSZ catalogue in terms of completeness, statistical reliability, and accuracy of cluster parameters including size and photometry. Section 4 extensively describes validation of cluster candidates through pre-existing surveys and cluster catalogues in many wavebands, while Sect. 5 describes the follow-up campaigns conducted by the *Planck* collaboration to confirm new cluster discoveries. This leads to a description of the catalogue properties in Sect. 6. The physical properties of the clusters are exploited in Sect. 7. These include an update of the SZ–X-ray scaling relations from the *Planck* data, the measure of the X-ray flux for all SZ detections, and the production of homogenized SZ-mass estimates for 813 clusters with measured redshifts that are provided to the community as a value-added element to the *Planck* SZ catalogue.

Throughout the article, the quantities M_{500} and R_{500} stand for the total mass and radius where the mean enclosed density is 500 times the critical density at the cluster redshift. The SZ flux is denoted Y_{500}, where $Y_{500}\,D_{\rm A}^2$ is the spherically-integrated Compton parameter within R_{500}, and $D_{\rm A}$ is the angular-diameter distance to the cluster. The physical cluster quantities are computed with a fiducial ΛCDM cosmology with $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\rm m} = 0.3$ and $\Omega_\Lambda = 0.7$. Furthermore, all the fits are undertaken in the log-log plane using the BCES orthogonal regression method of Akritas & Bershady (1996), with bootstrap resampling, which allows for intrinsic scatter as well as uncertainties in both variables. All uncertainties are given at 68 per cent confidence level and all dispersions are given in $\log_{10}$.

2. Construction of the *Planck* SZ Catalogue

2.1. Input Planck data

The *Planck* catalogue of SZ sources is constructed from the total intensity data taken during the first 15.5 months of *Planck* survey observations. Raw data were first processed to produce cleaned time-lines (time-ordered information) and associated flags correcting for different systematic effects; channel maps were then produced for all the observing frequencies (see details in Planck Collaboration VI 2013; Planck Collaboration II 2013). These maps, together with the associated beam characteristics, are the main inputs for the SZ-finder algorithms presented in Sect. 2.2. Following Planck Collaboration VIII (2011), we used the six highest-frequency *Planck* channel maps, from 100 to 857 GHz, to produce the catalogue of SZ detections. This optimizes the signal-to-noise (S/N) of the extracted SZ detections and the usable sky fraction; see Appendix A for the choice of channel maps.

In order to optimize the SZ detection, together with avoiding contamination of the PSZ catalogue by bright point sources (PS), the latter are masked from the channel maps prior to the SZ detection as detailed in the following. To construct the PS mask, we use the *Planck* Catalogue of Compact Sources (PCCS). The PCCS (Planck Collaboration XXVIII 2013) is a collection of single-frequency source catalogues, one for each of the nine *Planck* frequency channels. The six single *Planck*-HFI frequency PS catalogues are used to first produce individual-frequency masks constructed by masking a radius equivalent to 1.28 FWHM ($3\,\sigma_{\rm beam}$) around every point source detected with $({\rm S/N})_{\rm PS} \geq 10$. Then a single common PS mask (see Fig. 2), which is the union of the six individual HFI-frequency channel masks, is constructed. It is applied to all six highest-frequency *Planck* channel-maps to mask the point sources prior to running the algorithms to detect SZ signal. The masked regions are filled using a harmonic in-painting method based on that of Bajkova (2005), which has the advantage of eliminating the discontinuities caused by the masking. In order to avoid any possible artificial spurious detections at the edges of the in-painted area, we further reject detections within an expanded common mask, constructed using the same procedure as described above, but using a masking radius equivalent to 2.13 FWHM ($5\,\sigma_{\rm beam}$) and covering less than 2.9% of the sky.

Bright radio sources are known to exist at the centre of galaxy clusters, but they generally have steep spectra and hence their flux is significantly reduced at the six highest *Planck* frequencies where the PS mask is constructed and where the clusters are detected. The Perseus cluster (see Fig. 19 later and the associated discussion) is one exception, with a point source that is so bright that the cluster is masked and thus not included in the *Planck* SZ catalogue.

2.2. Detection Methods

The catalogue of SZ sources is the result of a blind multi-frequency search, i.e., no prior positional information on known clusters is used as input to the detection, by three detection algorithms briefly described below. These algorithms were described and tested using simulations (Melin et al. 2012). They were used to construct the Early SZ (ESZ) *Planck* sample by Planck Collaboration VIII (2011). All three assume priors on the cluster spectral and spatial characteristics, which optimize the SZ detection by enhancing the SZ contrast over a set of observations containing contaminating signals. In the following we present the cluster model used as a template by the SZ-finder algorithms and we briefly describe the three detection methods (for details we refer the reader to Herranz et al. 2002; Melin et al. 2006; Carvalho et al. 2009, 2011; Melin et al. 2012).

2.2.1. Cluster model

The baseline pressure profile model used in the detection methods is the generalized NFW (Navarro et al. 1997) profile of Arnaud et al. (2010). This profile model was constructed by combining the observed, scaled, X-ray pressure profile of 31 clusters from the REXCESS sample (Böhringer et al. 2007) for $R < R_{500}$[2], with the mean pressure profile from three sets of numerical simulations (Borgani et al. 2004; Nagai et al. 2007; Piffaretti & Valdarnini 2008) for $R_{500} < R < 5\,R_{500}$. New observational constraints on the pressure distribution at $R > R_{500}$ have become available. Planck Collaboration Int. V (2013) constrained the detection of the thermal pressure distribution out to about $3\,R_{500}$ through stacking of the observed SZ profiles of 62 nearby massive clusters detected with high significance in the *Planck* ESZ sample. The resulting profile is in agreement with that derived for the Coma cluster (Planck Collaboration Int. X

[2] R_{500} relates to the characteristic cluster scale $R_{\rm s}$ through the NFW concentration parameter $c_{500} = 1.177$ for the baseline profile ($R_{\rm s} = R_{500}/c_{500}$).

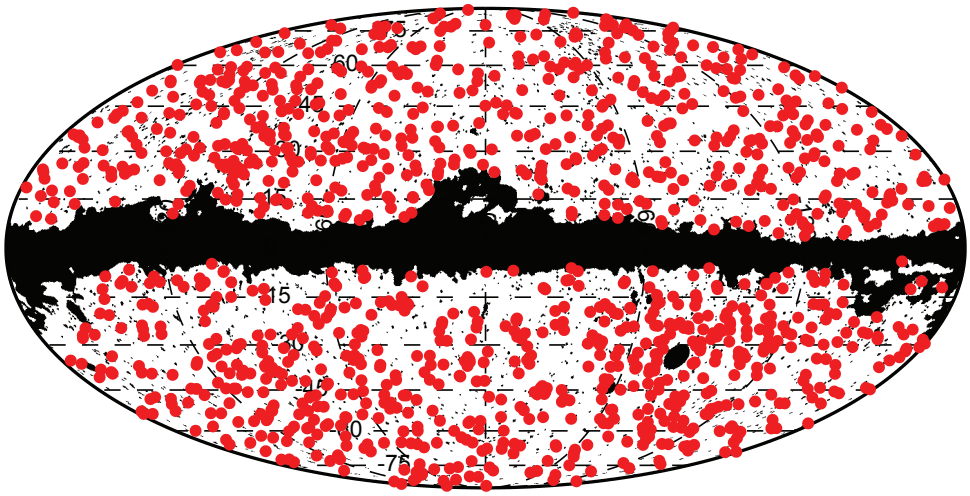


Fig. 2: The distribution, shown in Mollweide projection with the Galactic plane horizontal and the Milky Way centre in the middle, of the 1227 *Planck* clusters and candidates across the sky (red thick dots). The masked point-sources (black thin dots), the Magellanic clouds (large black areas) and the Galactic mask, covering a total of 16.3% of the sky and used by the SZ-finder algorithms to detect SZ sources, are also shown.

2013). Both show a slightly flatter distribution in the outer parts (i.e., beyond R_{500}) with respect to the predictions from the numerical simulations. These results are further confirmed by independent measurements from Bolocam in a smaller radial range ($r < 2\,R_{500}$, Sayers et al. 2012b). Using the profile of Planck Collaboration Int. V (2013) does not affect the detection yield (see Sect. 3) and only slightly modifies the measure of the SZ flux density (see Sect. 7.5) as compared to the generalized NFW (GNFW) profile adopted in the three cluster. The fiducial model parameters for the GNFW profile are given by the parameterization of the pressure profile in Eq. 12 of Arnaud et al. (2010). It states

$$\mathbf{p}(x) = \frac{P_0}{(c_{500}x)^{\gamma}\,[1 + (c_{500}x)^{\alpha}]^{(\beta-\gamma)/\alpha}}\,, \qquad (1)$$

with the parameters

$$[P_0, c_{500}, \gamma, \alpha, \beta] = [8.40\,h_{70}^{-3/2}, 1.18, 0.308, 1.05, 5.49]. \qquad (2)$$

The (weak) mass dependence of the profiles is neglected. Within the SZ-finder algorithms, the size and amplitude of the profile are allowed to vary but all other parameters are fixed. The cluster model is thus equivalent to a shape function characterized by two free parameters, its amplitude and a characteristic scale $\theta_{\rm s} = \theta_{500}/c_{500}$.

2.2.2. Matched Multi-filter (MMF)

Two different implementations of the matched multi-frequency filter algorithm (`MMF1` and `MMF3`) are used to detect SZ clusters. Both are extensions, over the whole sky, of the `MMF` algorithm (Herranz et al. 2002; Melin et al. 2006). The matched filter optimizes the cluster detection using a linear combination of maps (which requires an estimate of the statistics of the contamination) and uses spatial filtering to suppress both foregrounds and noise (making use of the prior knowledge of the cluster pressure profile and thermal SZ spectrum).

The `MMF1` algorithm divides the full-sky *Planck* frequency maps into 640 patches, each 14.66×14.66 square degrees, covering 3.33 times the sky. The `MMF3` algorithm divides the maps into a smaller set of 504 overlapping square patches of area 10×10 square degrees with the sky covered 1.22 times. The smaller redundancy of `MMF3` with respect to `MMF1` implies a potentially lower reliability of the SZ detections. In order to increase the reliability of the detections, the `MMF3` algorithm is thus run in two iterations. After a first detection of the SZ candidates, a subsequent run centred on the positions of the candidates refines the estimated S/N and candidate properties. If the S/N of a detection falls below the threshold at the second iteration, it is removed from the catalogue. For both implementations, the matched multi-frequency filter optimally combines the six frequencies of each patch. Auto- and cross-power spectra are directly estimated from the data and are thus adapted to the local instrumental noise and astrophysical contamination, which constitutes the dominant noise contribution. Figure 3 illustrates, for a six arcmin filter size, the ensemble noise maps as measured by `MMF3` in each of the patches. For both `MMF1` and `MMF3`, the detection of the SZ-candidates is performed on all the patches, and the resultant sub-catalogues are merged together to produce a single SZ-candidate catalogue per method.

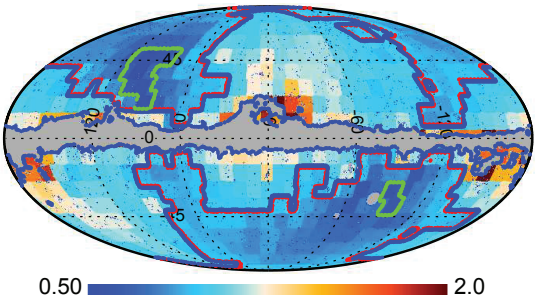


Fig. 3: The noise maps per detection patch of MMF3 method measured for a six arcmin filter. The noise ranges from 0.5 to 2 times the average noise of the map, which is $\sigma_Y = 2.4 \times 10^{-4}$ arcmin^2. The Ecliptic polar regions, delimited by green contours, with increased redundancy in the observations define a deep survey zone covering in total 2.7% of the sky. It is less noisy than the areas near the Galactic plane, where the dust emission is higher. Two other zones are defined: a medium-deep survey zone of 41.3% coverage delimited by the red contours and with higher noise level; and a shallow-survey zone covering 56% of the sky and with the highest noise levels including regions near the Galactic plane.

The candidate size in both algorithms is estimated by filtering the patches over the range of potential scales, and finding the scale that maximizes the S/N of the detected candidate. When merging the sub-catalogues produced from the analysis of individual patches, it is also the S/N of the detection (the refined S/N estimate for MMF3) which is used when deciding which detection of the candidate is kept. Furthermore, both MMF1 and MMF3 can also be run with fixed cluster size and position to estimate the SZ signal. This version of the algorithms is used to assess the reliability of the association with known clusters and/or to refine the measurement of the integrated Compton parameters of known X-ray clusters, as presented in Sect. 7.2.1.

2.2.3. PowellSnakes

PowellSnakes (PwS) is different from the MMF methods. It is a fast Bayesian multi-frequency detection algorithm designed to identify and characterize compact objects buried in a diffuse background. The detection process is grounded in a statistical model comparison test. The statistical foundations of PwS are described in Carvalho et al. (2009), and more recently in Carvalho et al. (2011) with a greater focus on the *Planck* setup. PwS may be run either based on a Generalized Likelihood Ratio Test or in full Bayesian mode. This duality allows PwS measured quantities to be consistently compared with those of the MMF algorithms.

PwS also operates in square flat patches of 14.66 × 14.66 square degrees. The total number of patches employed, of order 2800, varies with sky area but always guarantees a very large overlap; on average each cluster is detected about 4.7 times. PwS detects putative clusters and at the same time it computes the evidence ratio and samples from the posterior distributions of the cluster parameters. Then, it merges all intermediate sub-catalogues and applies the criterion of acceptance/rejection (Carvalho et al. 2011). PwS computes the cross-channel covariance matrix directly from the pixel data. To reduce the contamination of the background by the SZ signal itself, the estimation of the covariance matrix is performed iteratively. After an initial estimate, all detections in the patch with S/N higher than the current target detection are subtracted from the data using their best-fit values and the cross-channel covariance matrix is re-estimated. This is PwS '*native*' mode of background estimation that produces, on average, an S/N estimate about 20% higher than MMF. However, in order to produce a homogeneous *Planck* SZ catalogue from the three algorithms, it is possible to run PwS in '*compatibility*' mode, skipping the re-estimation step to mimic more closely the evaluation of the background noise cross-power spectrum of the MMF algorithms and thus their evaluation of the S/N. In this mode, PwS is a maximum likelihood estimator like the MMF.

In the following, unless stated otherwise, all quoted or plotted S/N values from PwS are obtained in '*compatibility*' mode in order to ensure homogeneity across the catalogue entries and in order to ease the comparison with the MMF outputs.

2.3. Outputs of the detection methods

Each of the three detection algorithms outputs a catalogue of SZ detections above S/N = 4.5 outside the highest-emitting Galactic regions (this corresponds to a mask of about 15% of the sky, see masked area in Fig. 2) and the Small and Large Magellanic Clouds and outside the PS mask described in Sect. 2.1. The union PS-Galactic mask covers 16.3% of the sky. The survey area used for the SZ detections in *Planck* is thus 83.7% of the sky coverage. The three individual lists of SZ candidates are cleaned by removal of obvious false detections. These are spurious sources that pass the MMF and PwS filters despite the pre-processing step applied to the *Planck* channel maps, see Sect. 2.1. In order to identify them, we cross-match the SZ detections with an intermediate, low signal-to-noise cut of 4, catalogue of point sources detected at the highest frequencies of *Planck*. Galactic sources in dense and cold regions at high latitudes also contaminate the SZ detections outside the Galactic mask. These cold Galactic sources (CGS hereafter, see Planck Collaboration XXIII 2011; Planck Collaboration XXII 2011) are detected in the *Planck* channel maps following an optimized method proposed by Montier et al. (2010). The SZ detections matching with PS at both 545 and 857 GHz, or with CGS sources, all show a rising spectrum at high frequencies, indicating that they are false detections. The SZ detections corresponding to such PCCS or CG sources are removed from the individual lists and from the published *Planck* catalogue of SZ sources.

The three detection algorithms used in the present study deploy the GNFW cluster profile to detect SZ signal with the two parameters of the shape function, the central value and the characteristic scale θ_s let free, with $\theta_s = \theta_{500}/c_{500}$. Each of the three algorithms therefore assigns, to each detected SZ candidate, a position with estimated uncertainty, a signal-to-noise value, and an estimated size, θ_s or equivalently θ_{500}, with its uncertainty. The detection likelihood or the posterior probability of the integrated Compton parameter within $5\theta_{500}$, denoted $Y_{5R_{500}}$, exhibits a large correlation with the size. Figure 4 illustrates the likelihood plots for two cases: a spatially-resolved cluster detected with a high signal-to-noise, Abell 2163; and a non-resolved cluster at high redshift ($z \simeq 1$), PSZ1 G266.6-27.3 (also known as PLCK G266.6-27.3 in Planck Collaboration XXVI 2011). We also show in Fig. 5 the distribution of maximum likelihood SZ fluxes ($Y_{5R_{500}}$) and sizes (θ_{500}) for the MMF3 detections.

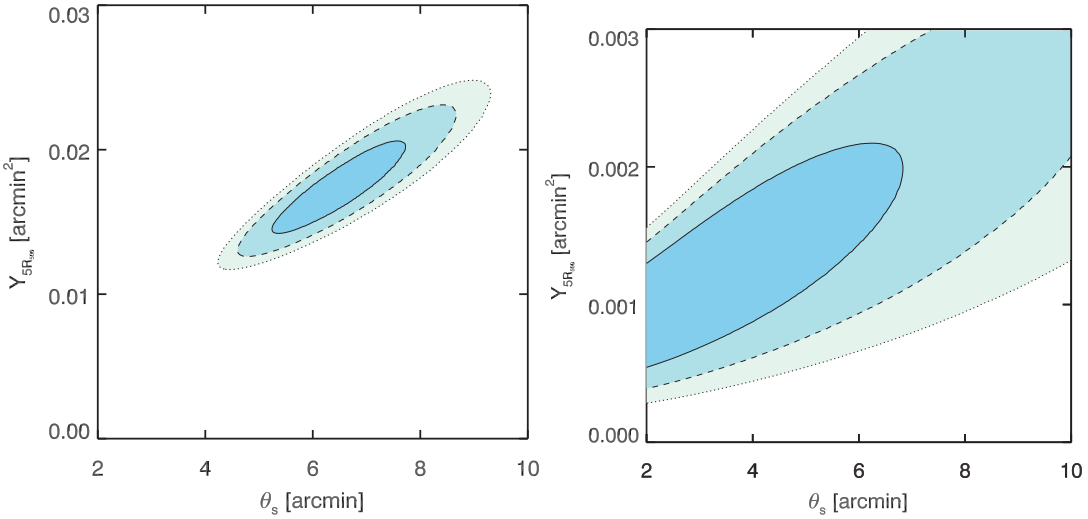


Fig. 4: Illustration of the SZ size–flux degeneracy for two clusters detected by *Planck*. *Right:* Abell 2163 (S/N = 27) and *left:* PSZ1 G266.6-27.3 (S/N = 6 at $z \simeq 1$). The contours show the 68, 95, and 99 percent confidence levels.

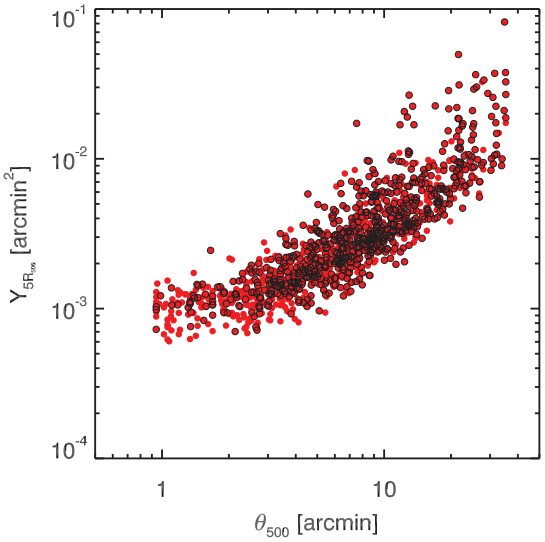


Fig. 5: Distribution of the maximum likelihood SZ flux $Y_{5R_{500}}$ and size θ_{500} for *Planck* SZ detections in the union catalogue down to S/N = 4.5. Detections associated with known or new confirmed clusters are shown as open black circles. SZ cluster candidates are shown as filled-red circles.

This "degeneracy" between cluster size and SZ flux propagates the size uncertainty to the SZ flux estimate, increasing and biasing its value dramatically. This effect being so detrimental, both the SZ blind flux and size best-fit estimates, and respective error bars, are not quoted in the catalogue outputs to avoid their misuse. Only the full joint $Y_{5R_{500}}-\theta_s$, or equivalently $Y_{5R_{500}}-\theta_{500}$, posterior probability contours provide a complete description of the information output by each detection method. They are thus provided for each detection. In order to use the flux measure, one ought to break the size–flux degeneracy. This can be achieved by a joint analysis with a high-resolution observation of the same objects, or by assuming a prior on, or fixing, the cluster size e.g., to the X-ray size. The SZ signal can then be re-extracted with an uncertainty much smaller than the variation of the joint Y–θ probability distribution.

We now perform a systematic comparison of the outputs of the three algorithms. We compare the S/N ratio. In addition and for purposes of illustration, we compare the best-fit blind Y value from maximum-likelihood or posterior probability outputs, namely $Y_{5R_{500}}$[3]. We show the comparison in Fig. 6, considering detections down to S/N = 4.5. We quantify the difference between a given quantity estimated by two different algorithms, Q_2 and Q_1, by fitting a power law to the data in the form $Q_2/Q_p = 10^A (Q_1/Q_p)^\alpha$ with a pivot $Q_p = 6$ for S/N and $Q_p = 4\times10^{-3}$arcmin^2 for $Y_{5R_{500}}$. The results are given in Table 1, including the scatter estimates. The raw scatter was estimated using the error-weighted vertical distances to the regression line. The intrinsic scatter on Y_{500} was computed from the quadratic difference between the raw scatter and that expected from the statistical uncertainties. Table 1 also lists the mean difference in logarithm, $\Delta(\log Q)=\log(Q_2/Q_1)$, computed taking into account both statistical errors and intrinsic scatter, estimated iteratively.

2.3.1. Signal-to-noise

A crucial ingredient of the SZ detection algorithms, either the MMFs or PwS, is the background cross-power spectrum used to estimate the noise level. It is evaluated from the data locally on a per-patch basis (see Fig. 3 for an example of the noise per patch across the sky). The algorithms, and implementations, slightly

[3] $Y_{5R_{500}}$ can be rescaled to Y_{500} for the fiducial GNFW model as $Y_{5R_{500}} = 1.79 \times Y_{500}$ (Arnaud et al. 2010).

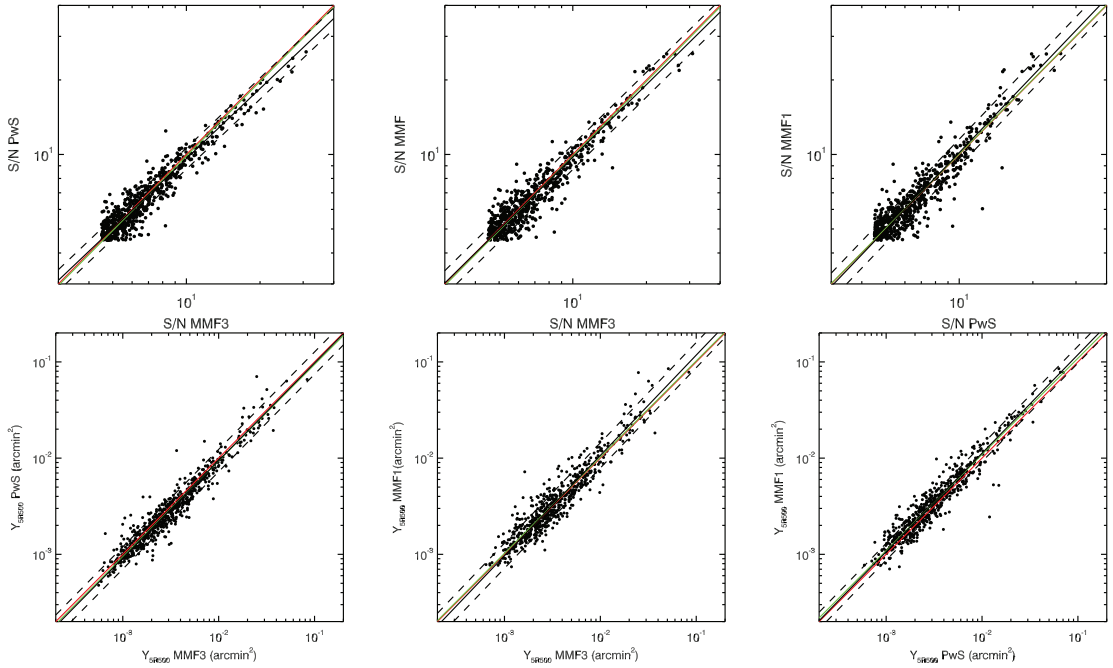


Fig. 6: Comparison of S/N (top panels) and maximum likelihood Compton-parameter values (bottom panels) from the three detection algorithms, MMFs and PwS, down to S/N = 4.5 after removing obvious false detections (see Sect. 2.3). In each panel, the red line denotes the equality line. The black line is the best fit to the data, and the dashed lines correspond to the $\pm 1\sigma$ dispersion about the fit relation. For clarity, error bars are omitted on $Y_{5R_{500}}$ values in the plot, but are taken into account in the fit. The green line of slope fixed to unity corresponds to the mean offset between the two quantities.

Table 1: Quantitative comparison between S/N and maximum likelihood Compton-parameter $Y_{5R_{500}}$ values from the three individual algorithms. Column(1): considered pair of algorithms ; Column(2-3): slope and normalization of the best-fit relation between the quantities estimated by the two algorithms, $Q_2/Q_p = 10^A\,(Q_1/Q_p)^\alpha$, using BCES orthogonal regression, with the pivot being $Q_p = 6$ for S/N and $Q_p = 4\times10^{-3}$ arcmin^2 for $Y_{5R_{500}}$; Column (4-5) intrinsic and raw scatter around the best-fit relation; Column (6): mean difference in logarithm, $\Delta(\log(Q))=\log(Q_2/Q_1)$; Column (7-8): corresponding intrinsic scatter and raw scatter.

	Power-law fit				Mean offset		
S/N	A	α	$\sigma_{\log}$		$\Delta\ \log(Q)$	$\sigma_{\log}$	
MMF3–PwS	-0.003 ± 0.002	0.94 ± 0.01	0.043 ± 0.002		-0.006 ± 0.002	0.045 ± 0.002	
MMF3–MMF1	-0.005 ± 0.002	0.97 ± 0.01	0.050 ± 0.002		-0.006 ± 0.002	0.051 ± 0.002	
PwS–MMF1	-0.000 ± 0.002	1.04 ± 0.02	0.054 ± 0.003		$+0.002 \pm 0.002$	0.054 ± 0.002	
$Y_{5R_{500}}$	A	α	$\sigma_{\log,\mathrm{int}}$	$\sigma_{\log,\mathrm{raw}}$	$\Delta\ \log(Q)$	$\sigma_{\log,\mathrm{int}}$	$\sigma_{\log,\mathrm{raw}}$
MMF3–PwS	-0.030 ± 0.004	1.01 ± 0.01	0.08 ± 0.03	0.116 ± 0.018	-0.027 ± 0.004	0.065 ± 0.006	0.102
MMF3–MMF1	$+0.011 \pm 0.005$	1.04 ± 0.02	0.11 ± 0.02	0.131 ± 0.014	$+0.010 \pm 0.005$	0.085 ± 0.006	0.118
PwS–MMF1	$+0.041 \pm 0.004$	1.02 ± 0.01	0.04 ± 0.01	0.088 ± 0.005	$+0.038 \pm 0.004$	0.040 ± 0.007	0.079

differ with respect to the stabilization assumptions (e.g., smoothing) of the background noise cross-power spectrum and to the treatment of the background SZ signal, now acting as a contaminant. These differences translate into variations in the S/N values per method. In particular, when operated in "*compatibility*" mode (without background cluster subtraction), PwS estimation of the background cross-power spectrum is more affected than the MMF by SZ signal contamination. The SZ signal adds an extra component to the background noise producing lower S/N estimates. This is particularly noticeable when the SZ signal is very strong compared with background (typically S/N $\geq$ 15).

Despite the differences in background estimates, the yields from the three algorithms agree. In the left panel of Fig. 7, we show that the detection counts as a function of S/N for each detection method are in good overall agreement. The right panel of Fig. 7 shows the fraction of common detections over the union of detections from all three algorithms as a function of S/N. Sources with S/N > 8.5 are detected by all three methods. However, we note that PwS number counts decrease more rapidly than MMF counts above S/N = 15. This reflects the behaviour of PwS in "*compatibility*" mode described above, which estimates a higher background than the MMF methods at high

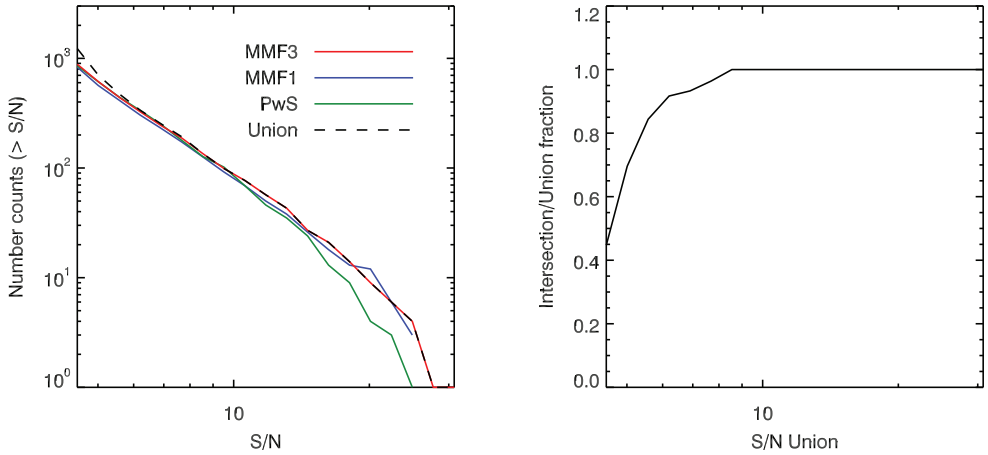


Fig. 7: *Left:* detection number counts as a function of S/N of the individual algorithms. The S/N value in the union catalogue is that of the MMF3 detections when available, followed by that of PwS followed by MMF1 (see Sect. 2.4). See text for discussion on the lower S/N values of PwS compared to the MMF-based algorithms. *Right:* fraction of common detection over counts from the union catalogue. Sources with S/N > 8.5 are detected by all methods.

S/N. Figure 6 shows the comparison of the S/N estimates from all three methods. The agreement is good on average. The mean ratio (or the normalization at the pivot of the power-law relation) deviates from unity by less than 2% and at less than $3\,\sigma$ significance. Here again at high S/N values, we note the tendency for lower S/N in PwS as compared to MMF (Fig. 6), and indeed the slope of the power-law relation is smaller than unity ($\alpha = 0.94 \pm 0.01$ for MMF3).

2.3.2. Photometry

We now compare the best-fit Y values (from maximum likelihood and posterior probability) for the three detection algorithms. The comparison (Fig. 6, lower panels) shows a systematic bias with PwS, yielding slightly smaller values than MMF, typically by 10%. However, the slope is consistent with unity, showing that this bias is not flux dependent. The MMF values differ from each other by less than 3% on average. The scatter between Y estimates is dominated by the intrinsic scatter (Table 1). It is clearly related to the size–flux degeneracy, the ratio between Y estimates for a given candidate being correlated with the size estimate ratio, as illustrated by Fig. 8. The scatter becomes compatible with the statistical scatter when a prior on the size is used, e.g., size fixed to the X-ray size.

2.4. Definition of the Planck SZ catalogue

As discussed above, the processing details of each algorithm/implementation differ in the computation of the background noise. The significance of the detections in terms of S/N, although in overall agreement, differs from one algorithm to the other and translates into different yields for the candidate lists from the three algorithms. We choose to construct a catalogue of SZ candidates that ensures, through redundant detections, an increased reliability of the low S/N sources, when they are detected by two methods at least, together with maximizing the yield of the catalogue.

The *Planck* SZ cluster catalogue described in the following is thus constructed from the union of the cleaned SZ-candidate lists produced at S/N ≥ 4.5 by all three algorithms. It contains in total 1227 SZ detections above S/N = 4.5. Note that in order to ensure homogeneity, in terms of detection significance, the S/N values of PwS quoted in the the union catalogue are obtained in *compatibility* mode, whereas the S/N obtained from PwS *native* mode are quoted in the PwS individual list. The union catalogue is constructed by merging detections from the three methods within an angular separation of at most five arcmin, in agreement with *Planck* position accuracy shown later in Fig. 12. As mentioned, no reference photometry is provided. However a reference position for the SZ detection is needed. For compatibility with the ESZ *Planck* sample, in the case of matching detection between methods we arbitrarily choose to take the coordinates from the MMF3 detection as the fiducial position (MMF3 was the reference method used to construct the ESZ *Planck* sample). When no detection by MMF3 above S/N = 4.5 is reported, we took the PwS coordinates as fiducial, and the MMF1 coordinates elsewhere. The S/N values in the union catalogue are taken following the same order, which explains why the MMF3 curve in Fig. 6 coincides with the union curve. The cluster candidates in the union catalogue are cross-referenced with the detections in the individual lists. The reference positions and the S/N values are reported in the union catalogue. Given the size–flux degeneracy, the full information on the degeneracy between size and flux is provided with each individual list in the form of the two-dimensional marginal probability distribution for each cluster candidate as discussed above. It is specified on a grid of 256×256 values in θ_s and $Y_{5R_{500}}$ centred at the best-fit values found by each algorithm for each SZ detection.

An extract of the *Planck* SZ catalogue is given in Appendix B. The full online table for union *Planck* catalogue, the individual lists of SZ detections, and the union mask used by the SZ-finder algorithms together with comments assembled

Table 2: Statistical characteristics of the *Planck* SZ catalogues. The intersection is defined as the set of SZ detections common between to three extraction algorithms. Completeness is a function of both Y_{500} and θ_{500}: the Y_{500} at a given completeness is estimated by marginalizing over θ_{500}, weighting each (Y_{500}, θ_{500}) bin by the theoretically-expected cluster counts. Positional accuracy is the median angular separation between real and estimated positions.

	Stat. Reliability(%)	$Y_{500}\ 10^{-3}$ arcmin^2 at **C**% completeness			Positional accuracy (arcmin)
		C=50	**C**=80	**C**=95	
Union	84	0.61	1.2	3.2	1.2
Intersection	98	0.85	1.8	6.6	1.1
MMF1	87	0.75	1.6	4.7	1.2
MMF3	91	0.71	1.5	3.8	1.2
PwS	92	0.65	1.4	3.2	0.9

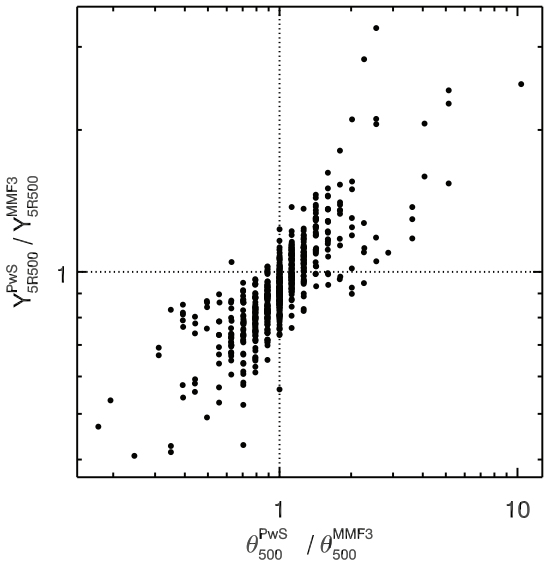


Fig. 8: Correlation between the ratio of $Y_{5R_{500}}$ estimates with PwS and MMF3 and the ratio of size estimates, shown on a grid of sizes.

in an external file are available at ESA's *Planck* Legacy Archive (PLA)[4].

3. Statistical Characterization

The statistical characterization of the PSZ catalogue is achieved through a process of MonteCarlo quality assessment (MCQA) that can be applied to each individual catalogue and to the merged union catalogues. The statistical quantities produced include completeness, fraction of detections associated with true clusters called, statistical reliability or purity, positional accuracy, and accuracy of parameter estimation. Together, these statistics describe the quality of detections in the catalogue. The quality of the parameter estimation, including astrometry (cluster position and extent), is determined through comparison with the parameters of the input clusters. The statistical characteristics of the different lists are summarized in Table 2.

3.1. MCQA Pipeline and simulations

The MCQA pipeline contains a common segment producing simulated input catalogues and processed, source-injected maps, which are then fed into the detection pipeline. In summary, the pipeline steps per MonteCarlo loop are:

1. creation of an input cluster catalogue;
2. injection of clusters into common simulated diffuse frequency maps, including beam convolution;
3. injection of multi-frequency point sources;
4. pre-processing of maps, including masking and filling point sources;
5. detection and construction of individual cluster-candidate catalogues;
6. construction of a union catalogue given merging criteria;
7. collation of input and output catalogues, producing detection truth-tables and catalogues of unmatched spurious detections[5].

To estimate the completeness, clusters are injected into the real data. In this case, steps 3 and 4 are skipped and each detection algorithm estimates noise statistics on the real data prior to injection in order to avoid artificially raising the S/N and biasing the completeness estimates. The pressure profiles of the injected clusters follow that described in Sect. 2.2.1. To account for the profile variation across the cluster population, the profile parameters are drawn from the covariance matrix of the 62 measured pressure profiles from Planck Collaboration Int. V (2013), ensuring that the injected profiles are consistent with measured dispersion and consistent, on average, with the extraction filter. The injected clusters are convolved with effective beams in each pixel including asymmetry computed following Mitra et al. (2011).

The simulated input cluster catalogues differ for statistical reliability and completeness determination. For completeness, clusters injected in real data are drawn from a uniform distribution in (Y_{500}, θ_{500}) so as to provide equal statistics in each completeness bin. To avoid an over-contamination of the signal, injected clusters are constrained to lie outside an exclusion radius of $5R_{500}$ around a cluster, either detected in the data or injected.

For the statistical reliability estimation of the input cluster distribution injected in simulations is such that cluster masses and redshifts are drawn from a Tinker et al. (2008) mass function and converted into the observable parameters (Y_{500}, θ_{500})

[4] http://www.sciops.esa.int/index.php?page=Planck_Legacy_Archive&project=planck.

[5] A cluster is considered to be matched if there is a detection within five arcmin of its position.

using the *Planck* ESZ Y_{500}–M_{500} scaling relation (Planck Collaboration X 2011). The simulated maps consist of CMB realizations, diffuse Galactic components and instrumental noise realizations, including realistic power spectra and inter-detector correlations, from the FFP6 simulations (Planck Collaboration XII 2013; Planck Collaboration ES 2013). Residual extragalactic point sources are included by injecting, mock-detecting, masking and filling realistic multi-frequency point sources using the same process as for the real data (see Sect. 2.1).

3.2. Completeness

The completeness is the probability that a cluster with given intrinsic parameters (Y_{500}, θ_{500}) is detected given a selection threshold (here in S/N).

If the Compton-Y estimates are subject to Gaussian errors, the probability of detection per cluster follows the error function and is parameterized by $\sigma_{Yi}(\theta_{500})$, the standard deviation of pixels in the multi-frequency matched-filtered maps for a given patch i at the scale θ_{500}, the *intrinsic* Compton Y_{500}, and the detection threshold q:

$$P\left(d|Y_{500}, \sigma_{Yi}(\theta_{500}), q\right) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{Y_{500} - q\sigma_{Yi}(\theta_{500})}{\sqrt{2}\sigma_{Yi}(\theta_{500})}\right)\right], \quad (3)$$

where $\mathrm{erf}(x) = (2/\pi)\int_0^x \exp\left(-t^2\right) dt$ and d is the Boolean detection state.

The completeness of the catalogue, thresholded at S/N q, is expected to follow the integrated per-patch error function completeness

$$C(Y_{500}, \theta_{500}) = \sum_i f_{\mathrm{sky},i} P\left(d|Y_{500}, \sigma_{Yi}(\theta_{500}), q\right), \quad (4)$$

where $f_{\mathrm{sky},i}$ is the fraction of the unmasked sky in the patch i. The true completeness departs from this theoretical limit. This is due to the non-Gaussian nature of the noise dominated by the astrophysical, namely Galactic, contamination. This is also the case when the actual cluster pressure profile deviates from the GNFW used in the SZ-finder algorithms, or when the effective beams deviate from constant symmetric Gaussians, and also when the detection algorithm includes extra steps of rejection of spurious sources not formulated in Eq. 3. This is why an MCQA-based assessment of the completeness is essential to characterize the *Planck* detections.

The MonteCarlo completeness of each of the individual lists and the union catalogue are shown in Fig. 9. The MMF lists are consistent with one another at $\theta_{500} > 4$ arcmin, but MMF3 is more complete at lower radii. This is due to an extra step implemented in MMF1 that rejects as spurious the detections estimated to be point-like. The union improves upon the completeness of each of the individual catalogues, because it includes the faint real detections by one method alone. In contrast, the intersection of the lists from the three algorithms, while more robust, is markedly less complete than the union and each of the individual catalogues. The intersection and union catalogues represent the extremes of the trade-off between statistical reliability and completeness. The quantities for each of the catalogues, plus the union and intersection, are summarized in Table 2. Figure 9 shows four constant θ_{500} slices through the completeness contours for MMF3, comparing the MCQA-based completeness with the integrated error function completeness. At radii smaller than 6 arcmin, the MCQA-based completeness is systematically less complete, and the drop-off of the completeness function shallower, than the theoretical expectation. This effect is a consequence of the variation of intrinsic cluster profiles from the GNFW profile assumed for extraction.

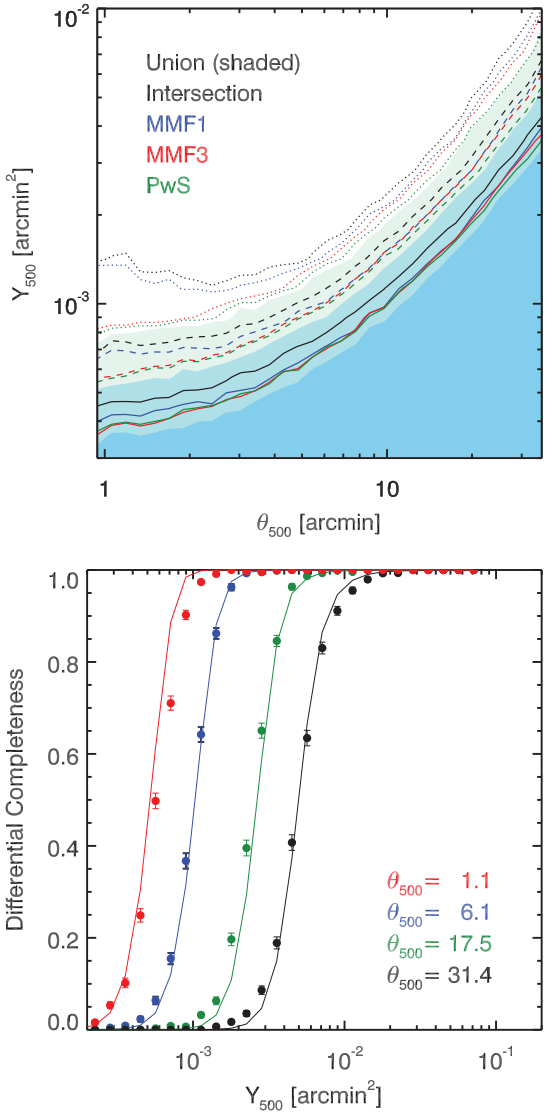


Fig. 9: *Top panel:* differential completeness as a function of (Y_{500}, θ_{500}) for each detection algorithm (MMF1 in blue, MMF3 in red, and PwS in green) and for the union (shaded area) and intersection (black) catalogues. From bottom to top, the solid, dashed, and dotted lines show 15%, 50% and 85% completeness, respectively. *Bottom panel:* slices through the MCQA-based completeness function at various θ_{500} for MMF3 compared to the error function approximation (solid curves).

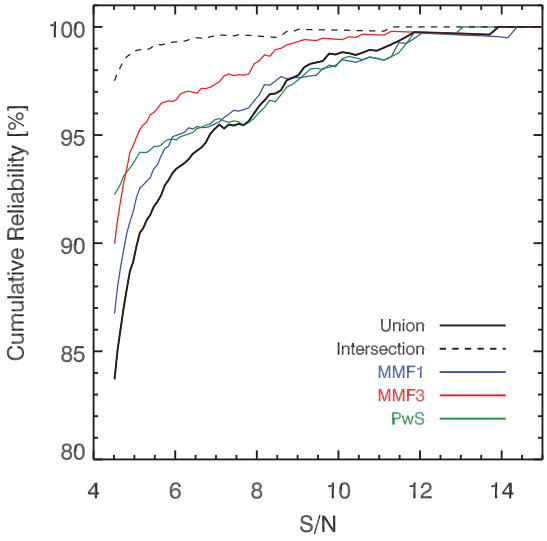


Fig. 10: Cumulative statistical reliability, defined as the fraction of sources above a given S/N associated with a "real" cluster from the simulated input catalogue.

3.3. Statistical reliability

The fraction of detections above a given S/N that are associated with a real cluster is characterized by injecting clusters into high-fidelity simulations of the *Planck* channels. Unassociated detections from these simulations define the fraction of spurious detections. We have verified that the simulations produced detection noise $\sigma_{Y_{500}}$ consistent with the real data and that the simulated detection counts match the real data.

The cumulative fraction of true clusters, as characterized by the simulations, is shown for the output of each detection algorithm and for the union catalogue in Fig. 10. The union catalogue is less pure than any of the individual lists because it includes all the lower-reliability, individual-list detections, in addition to the more robust detections made by all three SZ-finder algorithms. The union catalogue constructed over 83.7% of the sky at S/N of 4.5 is 84% pure.

The fraction of false detections is dominated by systematic foreground signals, in particular Galactic dust emission. This is illustrated in Fig. 11 by the effect of dust contamination on the cumulative reliability. We define two sky regions by the level of dust contamination: "region 1" is the low dust-contamination region outside of the *Planck* Galactic dust, and PS, mask that excludes 35% of the sky. This mask is used in Planck Collaboration XX (2013) for cosmological analysis of SZ counts. "Region 2" is the complementary region included by the smaller 15% dust mask but excluded by the 35% mask. When the larger Galactic dust mask is applied leaving 65% of the *Planck* sky survey in which to detect SZ signal, the statistical reliability increases from 84% in 83.7% of the sky to 88% in 65% of the sky. As seen in Fig. 11 upper panel, the reliability of the detections deteriorates markedly in "region 2" relative to "region 1". The noisy behaviour of the curves in Fig. 11 upper panel is due to the reduced size of sky area used in the analysis.

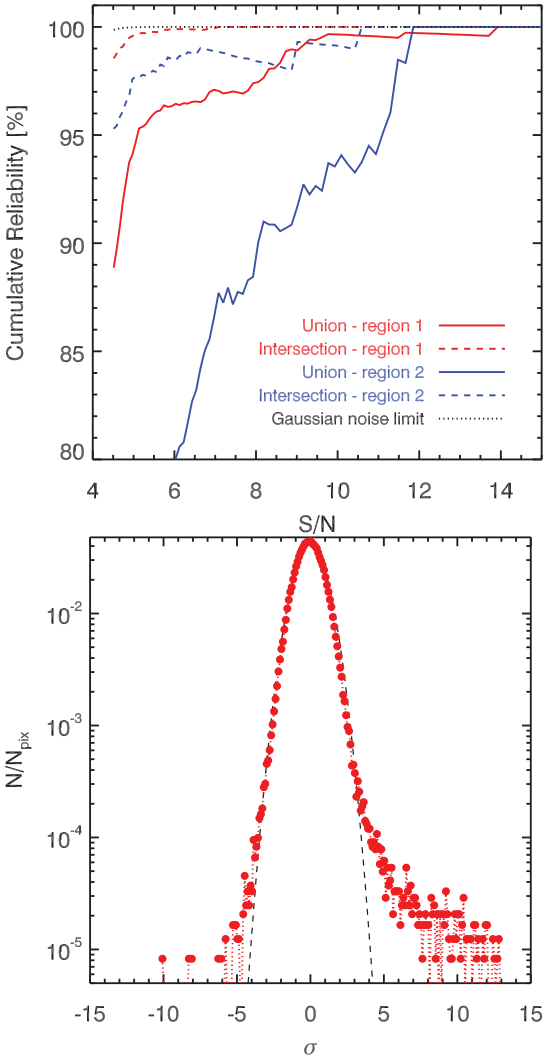


Fig. 11: *Top panel:* cumulative reliability for the union and intersection catalogues, as a function of dust contamination. Region 1 is the low-dust contamination region, being the 65% of the sky outside the Galactic dust mask, and region 2 is the complementary dustier region added to this when the smaller 15% dust mask is applied. The Gaussian noise limit is the expected reliability from purely Gaussian fluctuations. *Bottom panel:* histogram of the y-signal in a typical filtered patch from a null-test simulation, compared to the best-fit Gaussian (black dashed line). The distribution of y-noise is non-Gaussian.

In both regions, the spurious count much higher than is predicted by Gaussian fluctuations. This reflects the non-Gaussian nature of the filtered patches. The bottom panel of Fig. 11 illustrates this for a typical mid-latitude patch from a null-test simulation with no injected clusters. The patches are well approximated as Gaussian at deviations smaller than $3\,\sigma$ (consis-

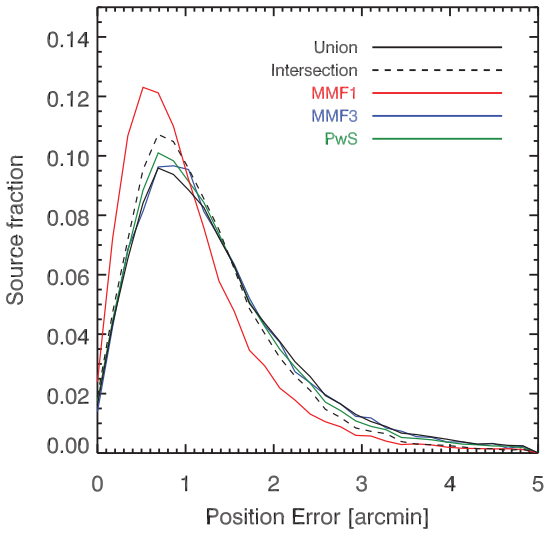


Fig. 12: Distributions of positional error for each catalogue, normalized by the total number of detections in the catalogue. By construction, the positional error is defined to be less than five arcmin.

tent with the assumptions of Eq. 3), but show enhanced numbers of high significance deviations, which can translate into spurious detections.

3.4. Positional Accuracy

Positional accuracy is characterized by the radial offset between estimated and injected positions. The distribution of position error is shown in Fig. 12, for each individual list and the union catalogue. In contrast to the MMFs, which estimate the maximum-likelihood position, the PwS position estimator is the mean of the position posterior, which produces more accurate positional constraints. The union catalogue positions are taken from MMF3 if available, followed by PwS and then MMF1. Its positional estimates are hence consistent with the MMFs. The mode of the union distribution is consistent with a characteristic position error scale of half an HFI map pixel (0.86 arcmin).

3.5. Parameter Recovery

The Compton $Y_{5R_{500}}$ is characterized by comparing detected and input values for matched detections from the injection of clusters into the real data (see Fig. 13). The injection follows the scheme outlined above with one exception: input cluster parameters are drawn using the Tinker mass function and the scaling relations discussed above for reliability simulations. This ensures a realistic distribution of parameters and S/N values.

What we characterize is slightly different for each catalogue. For the MMFs, we characterize the maximum-likelihood point of the 2-D degeneracy contours provided in the individual lists. For PwS, we characterize the mean of the marginal distribution for each parameter. In each case, the 2D ($Y_{5R_{500}}, \theta_s$) are marginalized over position. The contours are scaled for each cluster and are time consuming to compute, so we characterize the parameters from a lower-resolution grid that is better suited to Monte-Carlo analysis.[6]

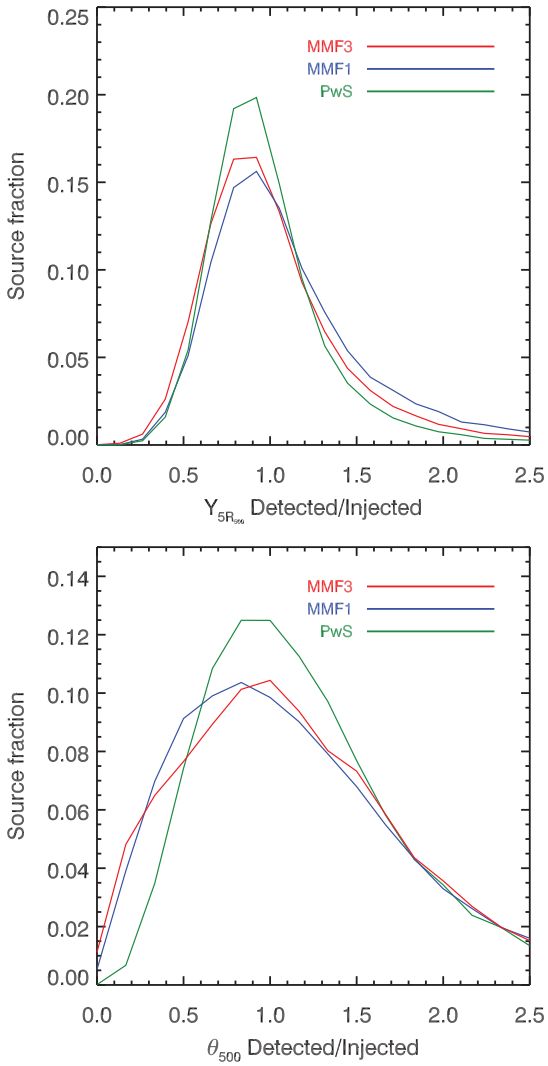


Fig. 13: Distributions of the ratio of detected over injected parameters for $Y_{5R_{500}}$ and θ_{500}.

The scatter between input and detected parameters is shown in Fig. 14 as an example for PwS. Biases are evident at both the low and high end for $Y_{5R_{500}}$. The low-flux bias is the Malmquist bias related to the S/N $\geq$ 4.5 threshold. The high-flux bias is due to a hard prior on the upper limit for cluster radius. Figure 14 also shows the distribution of the ratio of estimated over injected parameters. The median and median absolute deviation of these ratios are shown in Table 3.

[6] PwS does not resort to a low-resolution scale grid and always works at the full resolution.

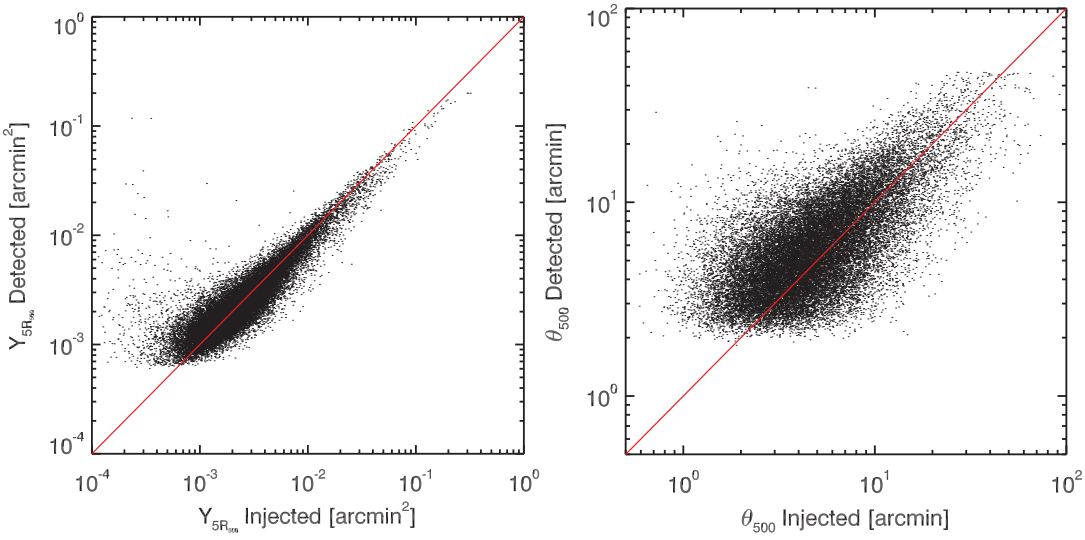


Fig. 14: Injected versus detected values of $Y_{5R_{500}}$ (left panel) and θ_{500} (right-panel), illustrated for PwS.

Table 3: Median and median absolute deviation (MAD) of the ratio of detected over injected parameters.

	$Y_{5R_{500}}$		θ_s	
	median	MAD	median	MAD
MMF1	1.09	0.39	1.17	0.70
MMF3	1.02	0.34	1.19	0.69
PwS	0.99	0.27	1.21	0.56

The distributions for flux are positively skewed due to Malmquist bias. The median ratios of the flux recoveries are consistent with unity for MMF3 and PwS and are slightly higher for MMF1. The recovery of θ_s is biased high in the median by about 20% for each of the codes. This bias is a consequence of the intrinsic cluster profile variation and disappears when the injected profiles match the detection filter. The $Y_{5R_{500}}$ estimate by contrast is relatively unaffected by profile variation. The parameter constraints from PwS are tighter than the MMFs due to the PwS priors and the definition of the estimator as the expected value of the parameters rather than the maximum likelihood.

4. External Validation

The cluster-candidate catalogue constructed from the union of all three SZ-finder algorithms undergoes a thorough validation process that permits us to identify previously-known clusters and to assess the reliability of the *Planck* SZ candidates not associated with known clusters. In order to achieve this, we make use of the existing cluster catalogues and we also search in optical, IR, and X-ray surveys for counter-parts at the position of the *Planck* SZ sources. In Sect. 5, we present the follow-up programmes that were undertaken by the *Planck* collaboration in order to confirm and measure the redshifts of the *Planck* candidate new-clusters.

The first step of the validation of the PSZ catalogue is to identify among the *Planck* SZ candidates those associated with known clusters. For this purpose, we use existing X-ray, optical or SZ cluster catalogues. A positional matching is not sufficient to decide on the association of a *Planck* SZ source with a previously-known cluster, and a consolidation of the association is needed. For the X-ray associations, a mass proxy can be built and used to estimate the SZ flux, S/N, etc, that are compared with measured quantities for the *Planck* cluster candidates. In contrast to the X-ray clusters, optical clusters either have no reliable mass estimates or suffer from large uncertainties in the mass–richness relations. In this case, the consolidation cannot be performed uniquely through the coherence of measured versus predicted properties. It rather relies on extra information from surveys in the X-ray, optical, or IR at the *Planck* cluster-candidate positions.

In the following, we detail the search for counter-parts in optical, IR, and X-ray surveys; list the cluster catalogues used for the identification; and finally present the identification procedure followed to associate *Planck* SZ detections with bona fide clusters. In this process, we define quality flags for the association of *Planck* SZ detections with external data. We set $Q = 1$ for high-reliability associations, i.e., very clear cluster signatures, $Q = 2$ for reliable associations, and $Q = 3$ for low-reliability associations, i.e., unclear cluster signature.

4.1. Search for counter-parts of Planck detections in surveys

We made use of the *ROSAT* All Sky Survey (RASS, Voges et al. 1999), the all-sky survey with the Wide-field Infrared Survey Explorer (WISE, Wright et al. 2010), and the Sloan Digital Sky Survey (SDSS, York et al. 2000) to search for counter-parts of the *Planck* SZ detections. This information was used in two ways. When *Planck* detections were associated with known clusters from catalogues, in particular in the optical, the counterparts in RASS, WISE, or SDSS helped in consolidating the asso-

ciation, increasing the confidence in the identification of *Planck* candidates with known clusters. When no association between *Planck* detections and previously-known clusters was found, the information on the counter-parts, in the surveys, of *Planck* SZ detections was used to assess the reliability of the *Planck* cluster candidates, i.e., clear or unclear cluster signatures.

4.1.1. Search in RASS data

As detailed in Planck Collaboration Int. IV (2013), the validation follow-up with *XMM-Newton* has shown the importance of the RASS data to assess the reliability of the *Planck* sources. In particular, Planck Collaboration Int. IV (2013) showed that a large fraction of *Planck* clusters are detectable in RASS maps, but this depends on the region of the sky and on the ratio Y_{500}/S_X which exhibits a large scatter (see later in Fig. 31 the case of the PSZ sources). We therefore exploit the RASS data to consolidate the identification with clusters from optical catalogues (see below Sect. 4.3.2) and to assess the reliability of the *Planck* SZ candidates.

We first perform a cross-match with the RASS bright source catalogue (BSC, Voges et al. 1999) and the faint source catalogue (FSC, Voges et al. 2000) within a five-arcmin radius of the position of each of the *Planck* SZ detections. We then perform a reanalysis of the RASS data following the methodology and prescriptions given by Böhringer et al. (2000, 2004) and Reiprich & Böhringer (2002). We compute count-rate growth curves in order to check for the extension of the signal. We estimate the source flux from both the growth curve (when adequate) and from a fixed five-arcmin aperture radius with respect to the surrounding background (after PS subtraction). We then derive the associated signal-to-noise in RASS, $(S/N)_{RASS}$. For this, we make use of the RASS hard-band, [0.5–2] keV, data that maximize the S/N of the detections. We furthermore computed the source density map of the BSC and FSC catalogues and the associated probability that a *Planck* cluster candidate will be associated with a B/FSC source within a radius of five arcmin. For the BSC, the probability of chance association is relatively low, with a median <1 %. As detailed in Planck Collaboration Int. IV (2013), the correspondence of a *Planck* SZ-candidate with a RASS-BSC source is a semi-certain association with a real cluster, whereas for the FSC catalogue the probability of chance association is larger, 5.2%.

We define a quality flag, Q_{RASS}, for the association of *Planck* candidates with RASS counter-parts using both the signal-to-noise in RASS and the association with B/FSC sources. This is of particular importance for the *Planck* candidate new clusters. Based on the results from Planck Collaboration Int. IV (2013), the quality of the association with RASS counter-parts is high, $Q_{RASS} = 1$, for *Planck* cluster candidates matching a RASS-BSC source or with $(S/N)_{RASS} \geq 2$. We find a total of 887 out of 1227 *Planck* SZ detections in this category, with mean and median signal-to-noise of 7.4 and 5.8, respectively. The quality is poor, $Q_{RASS} = 3$, for RASS counter-parts with $(S/N)_{RASS} < 0.5$ in regions of reasonable depth (quantified by the probability of chance association with FSC sources being larger than 2.5% (Planck Collaboration Int. IV 2013)).

4.1.2. Search in SDSS data

We performed a systematic search for counter-parts in the SDSS Data Release DR9 (SDSS-III Collaboration et al. 2012) at the position of all the *Planck* SZ detections. This was performed based on a cluster-finder algorithm developed by Fromenteau et al. (2013) to search for red galaxy over-densities in the SDSS galaxy catalogues.

For each associated counter-part within a five arcmin circle centred at the position of the *Planck* SZ detection, a quality criterion is defined on the basis of a fit to the luminosity function and the associated mass limit, and on the number of galaxies within five arcmin, N_{gal}, such that we have $Q_{SDSS,dat} = 1$, i.e., high quality, for cases where $N_{gal} \geq 40$ and for masses $M_{200} \geq 5.7 \times 10^{14}\,M_\odot$, $Q_{SDSS,dat} = 2$, i.e., good quality, for N_{gal} between 40 and 20 for masses between $1.5 \times 10^{14}\,M_\odot$ and $5.7 \times 10^{14}\,M_\odot$, and $Q_{SDSS,dat} = 3$ otherwise.

The cluster-finder algorithm outputs the position of the counter-part (Brightest Cluster Galaxy (BCG) and barycentre) and the estimated photometric redshift. When spectroscopic data are available for the brightest selected galaxy a spectroscopic redshift is also reported. The outputs of the cluster-finder algorithm are compared to those obtained by Li & White (2013) from different method based on the analysis of the full photometric-redshift probability distribution function (Cunha et al. 2009). In this approach, the position and redshift in the SDSS data that maximizes the S/N are considered as the best estimates for the counter-parts of the *Planck* SZ detections.

4.1.3. Search in WISE data

The Wide-field Infrared Survey Explorer (WISE, Wright et al. 2010) provides an all-sky survey at 3.4, 4.6, 12, and 22 μm (W1, W2, W3, W4) with an angular resolution of 6.1 to 12.0 arcsec in the four bands.

We search for counter-parts of the *Planck* SZ detections in the WISE source catalogue in two ways. On the one hand, we run an adaptive matched filter cluster finder developed by Aussel et al. (2013), similar to the one described by Kepner et al. (1999), using the cluster members' luminosity function of Lin et al. (2012). The background counts were determined from the neighbouring square degree in the vicinity of the *Planck* cluster candidate, excluding regions of fifteen arcmin centred on candidate positions. On the other hand, we use a method developed by Aghanim & Fromenteau (2013) based on a search for overdensities of bright ($W1 \leq 17$) and red ($W1-W2 > 0$) sources within a five-arcmin radius circle centred on the position of *Planck* detections with respect to a background computed in a fifteen-arcmin radius area.

Aghanim & Fromenteau (2013) find that a good-quality association between a *Planck* SZ-detection and a counter-part overdensity in WISE data is reached when there are at least ten galaxies above $2\,\sigma$ in the five-arcmin search region, and when the corresponding fraction of galaxies is at least 30% of the total number of galaxies retained in the fifteen-arcmin circle. Performing the search for counter-parts of an ensemble of random positions on the sky, we compute the purity of the detections, i.e., the probability of a *Planck* candidate having a real counter-part in the WISE data as opposed to a chance association. The quality criterion for the association between *Planck* detection and WISE overdensity is high, $Q_{WISE} = 1$, for a purity larger than 90%. When it lies between 90% and 80% the association of *Planck* SZ-detections and WISE overdensities is assigned a lower quality criterion $Q_{WISE} = 2$. We set the quality of the association to $Q_{WISE} = 3$, bad, when the purity is below 80%. We find 856 *Planck* SZ detections with high or good quality counter-parts in WISE data, including 658 $Q_{WISE} = 1$ detections.

4.1.4. DSS images

Finally for each *Planck* cluster candidate, the second Digitized Sky Survey[7] (DSS) database was queried for a field of 5 × 5 arcmin^2 centred at the position of the *Planck* SZ detections in the *r* and *ir* bands. The DSS images were used for visual inspection.[8] Clusters and rich groups out to $z \simeq 0.3$ to 0.4 can easily be identified in these plates as an obvious concentration of galaxies. This qualitative information was thus used: (i) to consolidate some identifications of *Planck* SZ detections with previously-known clusters; (ii) to optimize our strategy for the follow-up observations of *Planck* candidates (see Sect. 5); and (iii) to qualitatively assess the reliability or significance of the *Planck* SZ detections.

4.2. Cluster catalogues

We now present the ensemble of catalogues that were used to identify the *Planck* SZ detections with previously-known clusters. In the case of the *ROSAT*- and SDSS-based catalogues, we have used homogenized quantities, see below, that allowed us to perform the identification with comparable association criteria, which ensures homogeneity in the output results.

MCXC meta-catalogue – For the association of *Planck* SZ candidates with previously-known X-ray clusters, we use the Meta-Catalogue of X-ray detected Clusters of galaxies (MCXC, Piffaretti et al. 2011, and reference therein) constructed from the publicly available *ROSAT* All Sky Survey-based and serendipitous cluster catalogues, as well as the *Einstein* Medium Sensitivity Survey. For each cluster in the MCXC several properties are available, including the X-ray coordinates, redshift, identifiers, and standardized luminosity, $L_{\rm X,500}$, measured within R_{500}. The MCXC compilation includes only clusters with available redshift information (thus X-ray luminosity) in the original catalogues. We updated the MCXC, considering the first release of the REFLEX-II survey (Chon & Böhringer 2012), the third public release of clusters from the MACS sample (Mann & Ebeling 2012), individual MACS cluster publications and a systematic search in NED and SIMBAD for spectroscopic redshift for clusters without this information in the *ROSAT* catalogues. This yields an ensemble of 1789 clusters with z and $L_{\rm X,500}$ values, adding 20 MACS clusters, 21 REFLEX-II clusters and 5 SGP clusters to the MCXC. For these clusters, the expected Compton-parameter, $Y_{500}^{L_{\rm X}}$, and size, $\theta_{500}^{L_{\rm X}}$, are estimated combining the M_{500}–$L_{\rm X,500}$ relation of Pratt et al. (2009) and the M_{500}–Y_{500} relation given by Arnaud et al. (2010). The expected S/N ratio, $(\mathrm{S/N})^{L_{\rm X}}$, is computed taking into account the noise within $\theta_{500}^{L_{\rm X}}$ at the cluster location. We furthermore supplement the updated MCXC with 74 clusters from *ROSAT* catalogues without redshift information and 43 unpublished MACS clusters observed by *XMM-Newton* or *Chandra*. For these 117 objects, only centroid positions are available. Finally, we considered the published catalogues from *XMM-Newton* serendipitous cluster surveys with available redshifts, the XCS catalogue (Mehrtens et al. 2012), the 2XMMi/SDSS catalogue (Takey et al. 2011) and the XDCP catalogue (Fassbender et al. 2011). However, these catalogues mostly extend the MCXC to lower masses and only two *Planck* candidates were found to be associated with these new clusters.

[7] http://stdatu.stsci.edu/dss/.

[8] Images from the RASS, SDSS and WISE surveys at the position of the *Planck* SZ detections were also inspected.

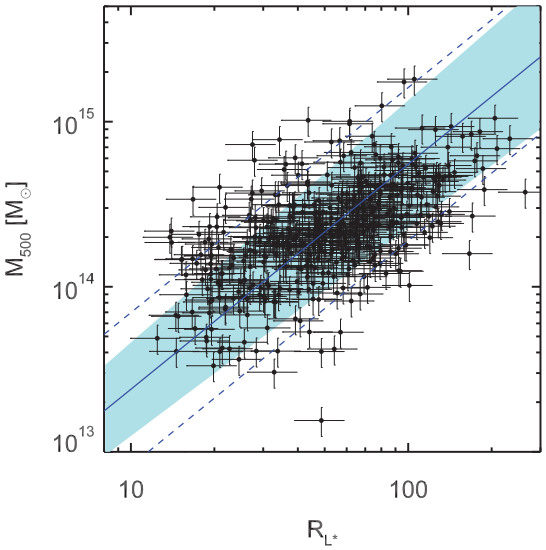


Fig. 15: Mass to richness scaling relation, M_{500}–$R_{\rm L^*}$, for the 444 MCXC clusters included in the WHL12 catalogue (Wen et al. 2012). The best-fit relation, from BCES fit, is given by the solid blue line. We adopted 15% uncertainties on the MCXC masses as prescribed in Piffaretti et al. (2011). As no uncertainty is provided for the WHL12's richness, we arbitrarily assumed a 20% uncertainty for all richness values. The blue shaded area shows the associated errors on the best-fit, while the dashed line marks the intrinsic scatter.

Optical-cluster catalogues – The identification of the *Planck* SZ candidates with clusters known in the optical is based on the Abell (Abell 1958) and the Zwicky (Zwicky et al. 1961) cluster catalogues. Furthermore, we have used four different catalogues of clusters based on the Sloan Digital Sky Survey (SDSS, York et al. 2000) data: (1) the MaxBCG catalogue (13,823 objects, Koester et al. 2007); (2) the GMBCG catalogue (55,424 objects, Hao et al. 2010); (3) the AMF catalogue (69,173 objects, Szabo et al. 2011); and (4) the WHL12 catalogue (132,684 objects, Wen et al. 2012). We refer the reader to Wen et al. (2012) for a comparison of the existing SDSS-based catalogues of clusters and groups. Each of the SDSS-based catalogues provides an estimated richness; we first start by homogenizing the richness estimates to that of WHL12. For each catalogue, we compute the median ratio of WHL12's richness to that of the considered catalogue over its intersection with WHL12's. We then renormalize the individual richness by the corresponding ratio. The correcting factors applied to the richness estimators[9] are respectively 1.52, 1.75, and 0.74 for MaxBCG, GMBCG, and AMF, obtained from 7627, 17245, and 1358 common clusters.[10] The richness is then related to the halo mass, M_{500}, by extending the Wen et al. (2012) richness–mass relation provided on about 40 clusters[11] to

[9] Field NGALS_R200 for MaxBCG, GM_SCALED_NGALS for GMBCG and LAM200 for AMF.

[10] We considered the associations of clusters with positions matching within 6 arcsec radius and with $\Delta z \leq 0.05$ (typical uncertainty for photometric redshifts in SDSS).

[11] Their M_{200} are taken from the literature either from weak lensing or X-ray measurements (Wen et al. 2010).

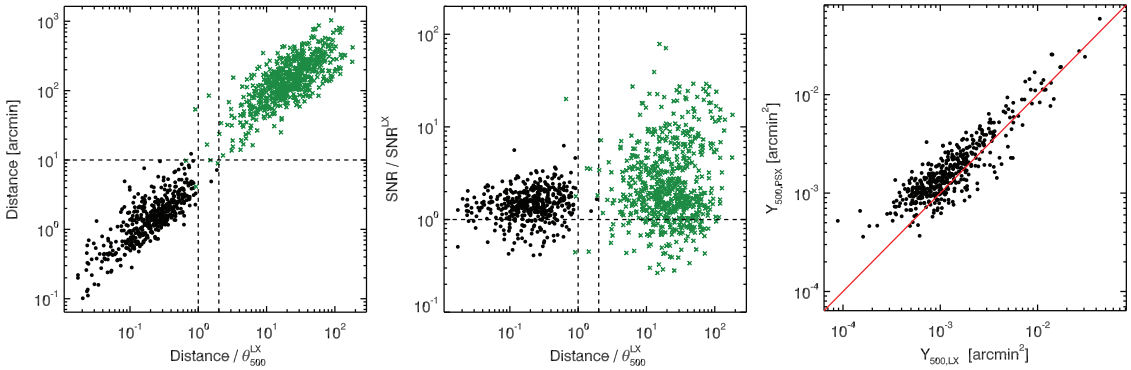


Fig. 16: Identification of the *Planck* cluster candidates with X-ray clusters from the MCXC catalogue. Black points are candidates firmly identified with MCXC clusters, while green points are candidates with no association. *Left panel:* distance of the *Planck* position to the position of the closest MCXC cluster as a function of the distance normalized to the cluster size $\theta_{500}^{L_X}$. *Middle panel:* S/N normalized to the expected value as a function of normalized distance. *Right panel:* SZ flux, $Y_{500,PSX}$, re-extracted fixing the position and size to the X-ray value, as a function of expected values. The red line is the equality line. In all panels, $Y_{500}^{L_X}$, and $\theta_{500}^{L_X}$ are estimated from the cluster X-ray luminosity used as mass proxy (see text).

444 MCXC clusters, with masses estimated from the X-ray luminosities. The data points and the best-fit scaling relation are presented in Fig. 15. The derived M_{500}–$R_{L\star}$ and $L_{X,200}$–$R_{L\star}$ relations are compatible with the findings of Wen et al. (2012). We find $\log(M_{500}/10^{14}\,M_\odot) = (-2.00\pm0.17)+(1.37\pm0.10)\times\log R_{L\star}$. The relation presents a large intrinsic log-scatter, $\sigma_{int} = 0.27 \pm 0.02$, hampering any accurate estimation of the cluster mass. This is further illustrated by the richest clusters with $R_{L\star} > 110$ having MCXC masses systematically below the best-fit M_{200}–$R_{L\star}$ relation (although within the 1σ intrinsic scatter).

SZ catalogues – At millimetre wavelengths, we cross-check the *Planck* SZ catalogue with the recent ACT and SPT samples (Menanteau et al. 2010; Vanderlinde et al. 2010; Williamson et al. 2011), including the most recent data that increased the number of SZ detections and updated the redshift estimates for the clusters (Reichardt et al. 2013; Hasselfield et al. 2013). We have furthermore identified the *Planck* SZ detections associated with previous SZ observations of galaxy clusters from the literature. We used a compilation of SZ observations conducted with the numerous experiments developed during the last 30 years (Ryle, OVRO, BIMA, MITO, Nobeyama, SZA, APEX-SZ, AMI, Diabolo, Suzie, Ryle, AMIBA, ACBAR, etc.).

4.3. Identification with previously-known clusters

4.3.1. Identification with X-ray clusters

The *Planck* SZ candidates are cross-checked against previously-known X-ray clusters from the updated version of the MCXC. For a given *Planck* candidate-cluster we identify the closest MCXC cluster.[12] The reliability of the association is assessed based on distance, D, compared to the cluster size and on the measured Y_{500} and S/N values compared with the expected values (see Fig. 16). Two clouds of points stand out in the scatter plot of absolute versus relative distance, $D/\theta_{500}^{L_X}$ (Fig. 16, left panel). They correspond to two clouds in the scatter plot of the measured over expected S/N ratio versus $D/\theta_{500}^{L_X}$ (Fig. 16, middle panel).

The association process follows three main steps. First, we provisionally assign an X-ray identification flag based on distance:

- $Q_X = 3$ if $D > 2\theta_{500}^{L_X}$ and $D > 10$ arcmin. Those are considered as definitively not associated with an MCXC cluster in view of *Planck* positional accuracy and cluster extent.
- $Q_X = 1$ if $D < \theta_{500}^{L_X}$ and $D < 10$ arcmin. Those are associated with an MCXC cluster.
- $Q_X = 2$ otherwise, corresponding to uncertain associations.

We then refine the classification. In the $Q_X = 1$ category, we identify outliers in terms of the ratio of measured to expected S/N and Y_{500}, taking into account the scatter and the size–flux degeneracy. Their flags are changed to $Q_X = 2$. In some cases, two distinct $Q_X > 1$ candidates are associated with the same MCXC cluster. The lowest S/N detection is flagged as $Q_X = 2$.

In the final step, we consolidate the status of $Q_X < 3$ candidates. We first re-extract the SZ signal at the X-ray position, both leaving the size free and fixing it at the X-ray value. The Y_{500} obtained with the cluster and size fixed to the X-ray values are compared to the expected values, $Y_{500}^{L_X}$, in the right panel of Fig. 16. For bona fide association, we expect no major change of Y_{500} and S/N, with, on average, a better agreement with the expected Y_{500} value and some decrease of S/N.

- For $Q_X = 1$ candidates, the re-extracted Y_{500} and S/N values are compared to both blind and expected values (as a function of distance, S/N, etc.) to identify potential problematic cases, e.g., important decrease of S/N or outliers in terms of measured-over-expected Y_{500} ratio. We found only one such case, whose flag is changed to $Q_X = 2$. The identification of other candidates is considered as consolidated, with definitive flag $Q_X = 1$.
- We then examine the $Q_X = 2$ candidates. We consider the re-extracted Y_{500} and S/N, but also perform a visual inspection of the SZ maps and spectra and ancillary data, including RASS and DSS images. The $Q_X = 2$ candidates were identified as clearly identified as multiple detections of extended

[12] The information of the second closest is also kept to identify potential confusion or duplicate associations.

clusters or duplicate detections of the same clusters by different methods that were not merged (the former are flagged as false detections, the latter are merged with the corresponding candidate in the union catalogue) or not associated (e.g., SZ sources clearly distinct from the MCXC clusters with no significant re-extracted signal at the cluster position and size).

Finally, for MCXC clusters without redshift and luminosity information, the association was only based on distance, setting $D_X < 5$ arcmin, and the consolidated based on visual inspection of SZ, RASS and DSS images and other ancillary information. Two cases were found to be a mis-identification. The SZ candidate was closer by chance to a faint XCS cluster, in the vicinity of the real counter-part (another MCXC cluster and an Abell cluster, respectively).

4.3.2. Identification with optical clusters

The *Planck* SZ candidates are associated with known clusters from optical catalogues (Abell, Zwicky, SDSS-based catalogues) on the basis of distance with a positional matching within a search radius set to five arcmin. The consolidation of the association was performed using the RASS information as described below, which allows us to mitigate the chance associations with poor optical galaxy groups and clusters.

SDSS-based catalogues – We have considered the four catalogues listed in Sect. 4.2. We define a quality criterion for the association, $Q_{\rm SDSS}$, in terms of cluster richness as a proxy of the cluster mass (see for instance Johnston et al. 2007; Rozo et al. 2009). We set the quality criterion, $Q_{\rm SDSS}$, to 3 for low reliability, to 2 for good reliability and to 1 for high reliability. The corresponding richness thresholds are 110 and 70 for $Q_{\rm SDSS} = 1/2/3$, respectively. The corresponding estimated masses (given the M_{500}–$R_{L^\star}$ relation) are $M_{500} > 6.5 \times 10^{14}\,{\rm M}_\odot$ and $M_{500} > 3.5 \times 10^{14}\,{\rm M}_\odot$. However due to the large scatter and associated uncertainty in the mass estimate from the mass–richness relation, we consolidate the association of the *Planck* candidates with SDSS clusters by combining the $Q_{\rm SDSS}$ with the RASS signal at the *Planck*-candidate position (see Sect. 4.1.1). In practice, only associations with $Q_{\rm SDSS} = 1$ or 2 and a signal-to-noise, measured at the *Planck* position in an aperture of five arcmin in the RASS survey, $({\rm S/N})_{\rm RASS} \geq 1$ are retained as firm identifications. We stress that our choice of richness thresholds is relatively conservative on average. Indeed, our $Q_{\rm SDSS} = 1$ and 2 matched candidates are found with high $({\rm S/N})_{\rm RASS}$ values as shown in Fig. 17, with mean $({\rm S/N})_{\rm RASS} = 7.1$ and 6.6 and median $({\rm S/N})_{\rm RASS} = 5.9$ and 5.4 for $Q_{\rm SDSS} = 1$ and 2 matches, respectively.

Abell and Zwicky catalogues – The *Planck* candidates are associated with Abell and Zwicky clusters on the basis of a positional matching within five arcmin. In the present case, we do not make use of any richness information in order to consolidate the association. We rather use here solely the RASS signal, $({\rm S/N})_{\rm RASS}$, at the SZ-candidate position. *Planck*-candidates associated with Abell or Zwicky clusters and with $({\rm S/N})_{\rm RASS} \geq 1$ are retained as firmly identified. For associations with $({\rm S/N})_{\rm RASS} < 1$, we decided on a firm identification only after checking the status of the counter-part in the WISE data and performing a visual inspection of the SZ signal and of the images from ancillary data, including DSS images.

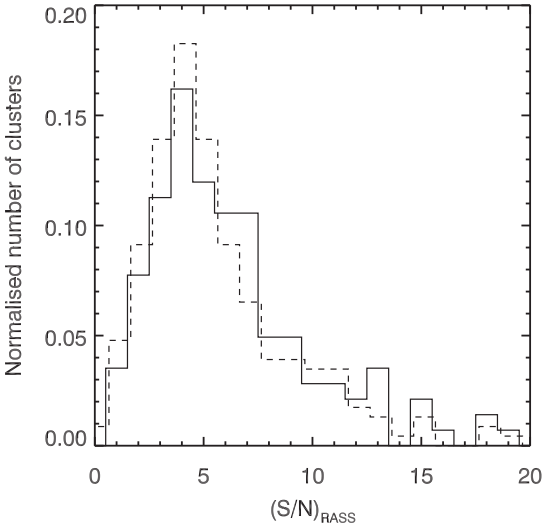


Fig. 17: Normalized distribution of the signal-to-noise in RASS survey at the position of *Planck* SZ detections with SDSS richness-based quality $Q_{\rm SDSS} = 1$ (solid line) and $Q_{\rm SDSS} = 2$ (dashed line).

4.3.3. Identification with SZ clusters

The association with known SZ clusters was performed within a five-arcmin radius. A visual inspection of the ancillary data and an *a posteriori* check of the RASS signal at the position of the *Planck* candidates associated with clusters from SZ catalogues is performed. It confirms that the values of $({\rm S/N})_{\rm RASS}$, when the coverage is significant, are high with an average value of 5.4.

4.3.4. Identifications from NED and SIMBAD

The information provided from querying NED and SIMBAD databases is mainly redundant with cross-checks with cluster catalogues. However, it lets us avoid missing a few associations. We therefore performed a systematic query in SIMBAD and NED with an adopted search radius set to five arcmin. Similarly to the association with clusters in optical catalogues, the positional association is consolidated using the results of the search in RASS data. Furthermore, the *Planck*-candidates solely matching NED or SIMBAD entries were inspected and the identification was confirmed or discarded using the information from WISE counter-parts and the DSS images.

5. Follow-up programme for confirmation of *Planck* candidates

We have undertaken, since Spring 2010, an extensive follow-up programme in order to perform a cluster-by-cluster confirmation of the *Planck* cluster candidates and obtain a measurement of their redshifts. A total of 276 *Planck* candidates, selected down to S/N = 4 from intermediate versions of the *Planck* SZ catalogue, were observed in pursuit of their redshift measurement. We have constructed our strategy for the selection of the *Planck* targets primarily on the successful results of the series of follow-up observations in X-rays based on Director's

Table 4: List of the main observing facilities used for the confirmation of the *Planck* newly-discovered clusters, and for the measurement of their redshifts.

Site	Telescope	Aperture (m)	Instrument	Filters	Redshift
...	XMM-Newton		EPIC/MOS & PN	...	Fe K
La Palma	NOT	2.56	ALFOSC	...	spec
La Palma	INT	2.5	WFC	*griz*	phot
La Palma	GTC	10.4	OSIRIS	...	spec
La Palma	TNG	3.5	DOLORES	...	spec
La Palma	WHT	4.2	ACAM	*griz*	phot
La Silla	NTT	3.7	EFOSC2	...	spec
La Silla	MPG/ESO 2.2m	2.2	WFI	VRI	phot
Mullard Radio Ast. Obs.	AMI	3.7 and 13	SA & LA	13.5 to 18 GHz	...
Tenerife	IAC80	0.82	CAMELOT	*griz*	phot
TUBITAK Nat. Obs.	RTT	1.5	TFOSC	*gri*	spec,phot

Discretionary Time on the *XMM-Newton* observatory (Planck Collaboration IX 2011; Planck Collaboration Int. I 2012; Planck Collaboration Int. IV 2013). Snapshot observations, sufficient to detect extended X-ray emission associated with *Planck* clusters and to estimate redshifts from the Fe line for the brightest clusters, were conducted sampling the SZ detections down to S/N = 4. These observations allowed us to better understand the SZ signal measured by *Planck* and hence to refine the criteria to select targets, especially for further optical follow-up.

We have engaged numerous campaigns on optical facilities, which now constitute our main means of confirmation of *Planck* SZ detections. *Planck* candidates with low-quality DSS images or without SDSS information, or low $(S/N)_{RASS}$, were primarily sent for deeper multi-band imaging observations. They were followed-up to the depth needed for the confirmation, i.e., finding an optical counter-part, and for the determination of a photometric redshift. Candidates with galaxy concentrations in DSS or with counter-parts in SDSS, and/or with high $(S/N)_{RASS}$, were preferentially sent for spectroscopic confirmation. The priority being to confirm the clusters and to secure the largest number of robust redshifts, no systematic spectroscopic confirmation of photometric redshifts was performed for low-redshift clusters ($z_{phot} < 0.4$). For higher-redshift clusters, spectroscopic confirmation of the photometric redshifts is more crucial. As a result, we have made use of telescopes of different sizes, from 1-m to 10-m class telescopes, optimizing the selection of targets sent to the different observatories (Table 4 gives the list of the main telescopes). Eight- and ten-metre class telescopes, e.g., GTC, GEMINI and VLT, were used to spectroscopically confirm redshifts above 0.5 for already confirmed clusters.

Our efforts to confirm the *Planck* cluster candidates, measure redshifts, and characterize cluster physical properties relies on ongoing follow-up of a large number of cluster candidates in the optical (ENO, RTT150, WFI), in the infrared (Spitzer[13]) and at SZ wavelengths (Arcminute Microkelvin Imager, AMI). The output of the confirmation and redshift measurements from the observing campaigns is summarized in Sect. 6.2. Companion publications, in preparation, will detail the observing campaigns and their results.

5.1. XMM-Newton observatory

The X-ray validation follow-up programme of 500 ks observations undertaken in *XMM-Newton* DDT is detailed in Planck Collaboration IX (2011), Planck Collaboration Int. I (2012), and Planck Collaboration Int. IV (2013). It consisted of observing 51 *Planck* targets and led to the confirmation of 43 *Planck* cluster candidates, two triple systems and four double systems. There were eight false candidates. This follow-up programme has constituted the backbone of the *Planck* cluster confirmation and most importantly has allowed us to better understand the SZ signal measured by *Planck* and thus to better master the criteria for confirmation (or pre-confirmation) of the *Planck* cluster candidates. By providing us with the physical properties and redshift estimates of the confirmed clusters, it has furthermore given us a first view on the physical characteristics of the newly discovered *Planck* clusters. Snapshot observations (around 10 ks) of the *Planck* candidates took place between May 2010 and October 2011. All the results from the four observing campaigns were published in Planck Collaboration IX (2011), Planck Collaboration Int. I (2012), and Planck Collaboration Int. IV (2013). Calibrated event lists were produced with v11.0 of *XMM-Newton*-SAS, and used to derive redshifts and global physical parameters for the confirmed clusters (Planck Collaboration IX 2011). The redshifts were estimated by fitting an absorbed redshifted thermal plasma model to the spectrum extracted within a circular region corresponding to the maximum X-ray detection significance. Most of the redshifts were confirmed using optical observations. Additional observations at VLT were conducted to confirm spectroscopically the highest redshifts.[14]

5.2. Optical observation in the Northern hemisphere

5.2.1. ENO telescopes

In total 64 cluster candidates from *Planck* were observed at European Northern Observatory (ENO[15]) telescopes, both for imaging (at IAC80, INT and WHT) and spectroscopy (at NOT, GTC, INT and TNG), between June 2010 and January 2013.[16] The aims of these observations were the confirmation, photo-

[13] Under Spitzer programs 80162 and 90233.

[14] Observations are conducted under programme 090A-0925.

[15] ENO: http://www.iac.es/eno.php?lang=en.

[16] The observations were obtained as part of proposals for the Spanish CAT time (semesters 2010A, 2010B, 2011A, 2011B, 2012A and 2012B), and an *International Time Programme (ITP)*, accepted by the International Scientific Committee of the Roque de los Muchachos (ORM, La Palma) and Teide (OT, Tenerife) observatories (reference ITP12_2).

metric redshift measurement, and spectroscopic confirmation of redshifts above $z = 0.3$.

INT, WHT and IAC80 – The optical imaging observations were taken either with the Wide-Field Camera (WFC) on the 2.5-m Isaac Newton Telescope (INT), the auxiliary-port camera (ACAM) at the 4.2-m William-Herschel Telescope (WHT), or with CAMELOT, the optical camera at the 0.82-m telescope (IAC80). The targets were observed in the Sloan *gri* filters. For the majority of fields, either Sloan *z* or Gunn *Z* images are also available. Images were reduced using the publicly-available software `Iraf` and `Sextractor` (Bertin & Arnouts 1996). The data reduction included all standard steps, i.e., bias and flat field corrections, astrometric and photometric calibrations. The photometric calibration is based either on standard star observations or, if available, on data from the SDSS. Finally, all magnitudes were corrected for interstellar extinction, based on the dust maps by Schlegel et al. (1998). We obtained photometric redshifts using the BPZ code (Benítez 2000), using a prior based on SDSS data, and fitting a set of galaxy templates. The BPZ code provides the Bayesian posterior probability distribution function for the redshift of each object, which is later used in the process of cluster identification. The identification of the galaxy overdensity located near the *Planck* positions and the estimate of the photometric redshifts of the associated clusters were performed using a modified version of the cluster-algorithm described in Sect. 4.1.2.

GTC and TNG – Spectroscopic observations were performed using the 10-m Gran Telescopio Canarias (GTC) telescope and the 3.6-m Telescopio Nazionale Galileo (TNG) telescope. The OSIRIS spectrograph at GTC was used in long-slit mode to observe a total of eight targets with two slit positions per candidate. We used the R500R grism and a binning 2×1, which provides a resolution $R = 300$ with a slit width 1 arcsec, and a wavelength coverage 4800–10000 Å. We retrieved three exposures of 1200 s each. The final spectra present a S/N of about 20 in galaxies with $r' = 20$ mag. We used the DOLORES multi-object spectrograph (MOS) at TNG to observe 9 candidates. The masks were designed to contain more than 30 slitlets, 1.5 arcsec width, placed within an area about 6 arcmin ×8 arcmin in order to cover the target field. We used the LR-B grism, which provides a dispersion of 2.7 Å/pixel, and a wavelength coverage between 4000 and 8000 Å. We carried out three acquisitions of 1800 s each and obtained spectra with S/N $\simeq$ 15 in galaxies with $r' = 20$ mag using a total integration time of 5400 s.

Nordic Optical Telescope (NOT) – Spectroscopic redshift measurements were obtained using the Andalucia Faint Object Spectrograph and Camera (ALFOSC) at the NOT.[17] Most targets were observed in MOS mode, targeting typically ten to fifteen galaxies per ALFOSC field (covering 6.4×6.4 arcmin^2, with an image scale of 0.188 arcsec/pixel). One or two unfiltered 300s pre-imaging exposures were obtained per candidate cluster, in addition to a single 300s exposure in each of the SDSS *g*-and *i* bands. The de-biased and flat field calibrated pre-imaging data were used to select spectroscopy targets. The final mask design[18] was carved out using custom software, generating slits of fixed width 1.5 arcsec and of length typically 15 arcsec. Grism No. 5 of ALFOSC was used, covering a wavelength range 5000 – 10250 Å with a resolution of about $R = 400$ and dispersion 3.1 Å/pixel. Redwards of 7200 Å strong fringing is present in the ALFOSC CCD. It was effectively suppressed using dither pattern alternating the placement of the spectroscopy targets between these sets of slits.

In addition to the MOS observations, spectroscopic observations in single-slit mode were conducted for some *Planck* candidates. For these observations, a long slit covering the entire 6.4 arcmin length of the ALFOSC field and a width of 1.3 arcsec was employed, with the same grism and wavelength coverage as for the MOS observations. The field angle was rotated to place the long slit over multiple targets, to include the apparent BCG as well as two to three other bright cluster galaxies within the ALFOSC field.

5.2.2. RTT150

A total of 88 *Planck* cluster candidates were followed up with the Russian Turkish Telescope (RTT150[19]) from July 2011 to December 2012 within the Russian quota of observational time. In total, about 50 dark nights, provided by Kazan Federal University and Space Research Institute (IKI, Moscow), were used for these observations. Direct images and spectroscopic redshift measurements were obtained using TÜBİTAK Faint Object Spectrograph and Camera (TFOSC[20]), similar in layout to ALFOSC at NOT (see above) and to other instruments of this series.

The TFOSC CCD detector cover a 13.3×13.3 arcmin^2 area with 0.39 arcsec per pixel image scale. Direct images of cluster candidates were obtained in Sloan *gri* filters, in series of 600s exposures with small ($\approx$ 10–30 arcsec) shifts of the telescope pointing direction between the exposures. All standard CCD calibrations were applied using `Iraf` software, individual images in each filter were then aligned and combined. The total of 1800 s exposure time in each filter was typically obtained for each field, longer exposures were used for more distant cluster candidates. Deep multi-filter observations were obtained for all candidates, except those unambiguously detected in SDSS. With these data, galaxy clusters can be efficiently identified at redshifts up to $z \approx 1$.

Galaxy clusters were identified as enhancements of surface number density of galaxies with similar colours. Cluster red sequences were then identified in the colour–magnitude diagram of galaxies near the optical centre of the identified cluster. The detected red sequence was used to identify the BCG and cluster member galaxies. Using the measured red-sequence colour photometric redshift estimates were obtained, which were initially calibrated using the data on optical photometry for galaxy clusters from the 400SD X-ray galaxy cluster survey (Burenin et al. 2007).

For spectroscopy we used the long-slit mode of the instrument with grism No. 15, which covers the 3900–9100 Å wavelength range with $\approx$ 12 Å resolution when a slit of 1.8 arcsec width is used. Galaxy redshifts were measured through the cross-correlation of obtained spectra with a template spectrum of an elliptical galaxy. Spectroscopic redshifts were typically ob-

[17] The observing runs took place on June 28 - July 3, 2011, January 20-25, 2012, July 16-21, 2012 and January 9-14, 2013.

[18] The MOS masks were cut at the Niels Bohr Institute, Copenhagen University.

[19] http://hea.iki.rssi.ru/rtt150/en/index.php.

[20] http://hea.iki.rssi.ru/rtt150/en/index.php?page=tfosc.

tained for the spectra of a few member galaxies, including the BCG, selected from their red sequence in the imaging observations. These data allow us to efficiently measure spectroscopic redshifts for clusters up to $z \approx 0.4$. For the highest-redshift clusters, complementary spectroscopic observations were performed with the BTA 6-m telescope of SAO RAS using SCORPIO focal reducer and spectrometer (Afanasiev & Moiseev 2005).

5.3. Optical observation in the Southern hemisphere

5.3.1. MPG/ESO 2.2-m Telescope

Optical imaging of 94 *Planck* cluster candidates in the Southern hemisphere was performed under MPG programmes at the MPG/ESO 2.2-m telescope using the Wide-Field Imager (WFI).[21] The WFI detector is a mosaic of 8 2k × 4k CCDs, covering a total area of 33 arcmin ×34 arcmin on the sky, with an image scale of 0.238 arcsec/pixel. Each field was observed in the V-, R-, and I-bands with a default exposure time of 1800s (with five dithered sub-exposures) per passband. The basic data calibration, including de-biasing and flat-field frame calibration, followed standard techniques. The individual exposures were re-registered and WCS calibrated using the USNO-B1 catalogue as an astrometric reference before being stacked into a combined frame for each filter, covering the entire WFI field. Photometric redshifts of the observed clusters were then determined from an algorithm that searches for a spatial galaxy overdensity located near the position of the SZ cluster candidate that also corresponds to an overdensity in $V - R$ versus $R - I$ colour–colour space. The median colour of galaxies located in this overdensity was then compared to predicted colours of early-type galaxies at different redshifts by convolving a redshifted elliptical galaxy spectral energy distribution template with the combined filter+telescope+detector response function.

5.3.2. New Technology Telescope (NTT)

Observations[22] were conducted at the 3.5-m NTT at the ESO observatory at La Silla to measure spectroscopic redshifts of 33 *Planck* clusters with the EFOSC2 instrument in the MOS mode. A clear BCG was identified in the clusters in pre-imaging data, and besides the BCG a redshift was measured for at least one other member of the cluster. In the following a brief outline of the observations and the data reduction are given (see Chon & Böhringer 2012, for details).

Each field of the *Planck* target candidates was optically imaged in Gunn r band for target selection and mask making. The imaging resolution is 0.12×0.12 arcsec^2, and the field of view is 4.1×4.1 arcmin^2 for both imaging and spectroscopic observations. When necessary, the field was rotated to optimize target selection. We used the grism that covers the wavelength range between 4085 Å and 7520 Å, with 1.68 Å per pixel at resolution 13.65 Å per arcsec. We typically applied 10 to 15 slitlets per field with a fixed width of 1.5 arcsec for the MOS and of 2.0 arcsec for the long-slit observations. Including at least three bright objects, preferably stars, to orient the field, the slitlets were allocated to the candidate member galaxies. The exposure times for the clusters range from 3600s to 10800s.

The data were reduced with the standard reduction pipeline of `Iraf`. The redshifts from the emission lines were determined separately after correlation with the passive galaxy templates. We use the `rvsao` package, which applies the cross-correlation technique to the input templates of galaxy spectra to measure the object redshift. The REFLEX templates were used for this analysis, which include 17 galaxy and stellar templates. We confirmed a spectroscopic cluster detection if at least three galaxies have their R-value greater than 5, and lie within ± 3000km/s of the mean velocity of the cluster members. We then took the median of those galaxy redshifts as the cluster redshift. For the long-slit observation, the cluster was confirmed with the redshift of the BCG and another galaxy at similar redshift within the aforementioned criteria.

5.4. Observations in the SZ domain with AMI

An ensemble of 60 *Planck* blind SZ candidates, spanning a range of S/N between 4 and 9 and meeting the Arcminute Microkelvin Imager (AMI) observability criteria, was observed with AMI. The goal of this programme was to confirm *Planck* cluster candidates through higher-resolution SZ measurements with AMI and to refine the position of confirmed clusters in order to optimize the subsequent optical follow-up observations aiming at redshift measurement. AMI comprises of two arrays: the Small Array (SA); and the Large Array (LA). Further details of the instrument are given in AMI Consortium et al. (2008). Observations carried out with the SA provide information that is well coupled to the angular scales of the SZ effect in clusters, whereas snapshot observations obtained with the LA provide information on the discrete radio-source environment. The latter allowed us to detect the presence of nearby, bright radio sources, helping in further selecting the targets for observation with the SA. Details of the AMI data reduction pipeline and mapping are described in Planck and AMI Collaborations (2013).

6. Results of the validation and follow-up

The external validation allows us to identify *Planck* SZ detections with previously-known clusters and to assemble crucial information on the identified clusters such as their redshifts. The validation steps corresponding to the association with known clusters were performed following a chosen hierarchy: X-ray clusters from the updated MCXC meta-catalogue; then optical clusters from Abell and Zwicky catalogues; then optical clusters from the SDSS-based catalogues; followed by SZ clusters from SPT and ACT samples; and finally clusters from NED and SIMBAD queries. The first identifiers of the *Planck* SZ detections given in Table B.1 reflect the validation hierarchy.

In the following, we present the results of the external validation process and of the follow-up campaigns for confirmation of *Planck* candidates and measurement of their redshifts (see Table 5 and Fig. 18). We also present the confirmation from SDSS galaxy catalogues and from X-ray archival data. We further discuss the unconfirmed candidate new clusters detected by *Planck*, which we classify into three categories of different reliability.

[21] Based on observations under MPG programmes 086.A-9001, 087.A- 9003, 088.A-9003, 089.A-9010, and 090.A-9010. The observations were conducted during the periods of November 27 - December 3, 2010, March 8-19, May 21 - June 3, and November 30 - December 4, 2011, December 30, 2011 - January 7, 2012, June 10-18, 2012, and January 6-13 2013.

[22] The observations were performed during three spectroscopic observing campaigns, 087.A-0740, 088.A-0268 and 089.A-0452.

Table 5: Summary of the external validation and confirmation from follow-up observations. In each category, the number represents the total number of *Planck* candidates identified exclusively with X-ray, optical, or SZ clusters. The category X-ray clusters covers identifications from the updated MCXC meta-catalogue. The category optical clusters covers identification from the Abell, Zwicky, and SDSS-based catalogues only. The category SZ clusters covers identification from SPT and ACT catalogues only. Confirmations from follow-up do not cover the observations performed by the *Planck* collaboration to measure the missing redshifts of known clusters. Confirmation from archival data covers X-ray data from Chandra, *XMM-Newton*, and *ROSAT* PSPC pointed observations only.

Planck SZ sources	1227		
Previously known clusters	683		
		X-ray only	472
		Optical only	182
		SZ only	16
		NED and SIMBAD	13
Planck new clusters	178		
		Confirmed with follow-up observations	157
		Confirmed with SDSS galaxy catalogues or archival data	21
Planck candidate new clusters	366		
		CLASS1	54
		CLASS2	170
		CLASS3	142

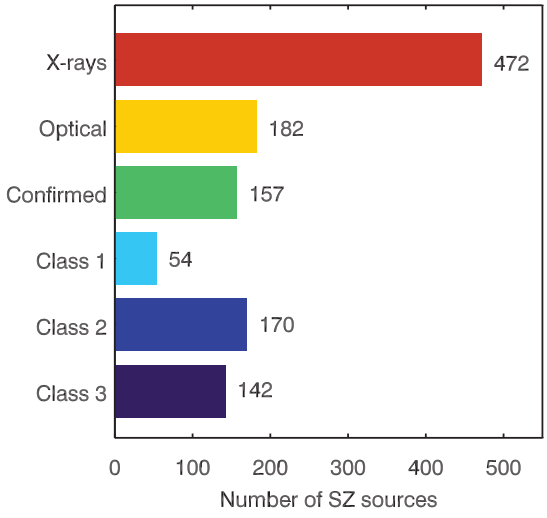


Fig. 18: Distribution of the *Planck* clusters and candidates in the different categories defined in the external validation process. The validation follows the order association with MCXC clusters, then Abell and Zwicky clusters, then SDSS clusters, then SZ clusters, and finally clusters from NED/SIMBAD.

6.1. Planck clusters associated with known clusters

A total of 683 out of 1227 SZ detections in the *Planck* catalogue, i.e., 55.7%, are associated with previously-known clusters from X-ray, optical, or SZ catalogues, or with clusters found in the NED or SIMBAD databases. We give the number of clusters identified in each category and we discuss notable cases of known clusters that are not included in the *Planck* SZ catalogue.

6.1.1. Identification with known X-ray clusters

A total of 472 *Planck* SZ-candidates are identified with known X-ray clusters from the MCXC meta-catalogue, which represents 38.5% of the *Planck* SZ detections and 69.1% of the identifications with previously-known clusters. These identifications of course account for many Abell clusters in the RASS-based catalogues of X-ray clusters.

Using the cluster properties reported in the MCXC and the *Planck* noise maps at the cluster positions, we computed the expected SZ signal and the expected S/N ratio for a measurement with *Planck*. We have compared the number of detected clusters in the *Planck* catalogue with S/N $\geq$ 4.5 to the number MCXC clusters at an expected significance of 4.5. Only 68 clusters expected to be detected at S/N > 4.5 are not included in the *Planck* catalogue, including 16 with predicted S/N between 4 and 4.5. Of the 52 clusters with expected S/N $\geq$ 4.5, only 41 are outside the masked regions and could thus be in the PSZ catalogue. Our computation of the expected SZ signal and S/N were based on scaling relations for X-ray-selected clusters, not accounting for the dispersion in the relations. We therefore focus on the non-detected MCXC clusters that significantly depart from the expected S/N value, namely by more than 5 σ. A total of 13 clusters are in this category. The two objects RXCJ2251.7-3206 and RXCJ0117.8-5455 show emission in high-resolution *Chandra* imaging that is point-like rather than extended and are likely not clusters of galaxies (Mantz et al. 2010; Magliocchetti & Brüggen 2007). Of the other eleven missing MCXC clusters, some present AGN contamination. This is the case for RXC J1326.2+1230 (Magliocchetti & Brüggen 2007), RXJ1532.9+3021 (Hlavacek-Larrondo et al. 2012), RXCJ1958.2-3011, RXCJ2251.7-3206, and RXCJ0117.8-5455 (Magliocchetti & Brüggen 2007), Abell

689 (Giles et al. 2012), ZwCl2089 (Rawle et al. 2012), PKS 0943-76 (Abdo et al. 2010), and Abell 2318 (Crawford et al. 1999). In these cases, the presence of the AGN affects the X-ray luminosity measure leading to an overprediction of the SZ signal. Some exhibit significant radio contamination, e.g., RXCJ1253.6-3931 (Plagge et al. 2010) and RXCJ1958.2-3011 (Magliocchetti & Brüggen 2007), which hampers the SZ detection. Cool-core clusters for which the X-ray luminosity is boosted due to the central density peak have an over-estimated expected SZ signal. This is the case for RXCJ0425.8-0833 (Hudson et al. 2010), ZwCl2701 (Rawle et al. 2012), Abell 1361 (Rafferty et al. 2008), and RBS 0540 (Eckert et al. 2011; Belsole et al. 2005). Other "missing" clusters are CIZA clusters: RXC J0643.4+4214, RXC J1925.3+3705, RXC J2042.1+2426 and RXC J0640.1-1253, REFLEX cluster RXCJ2149.9-1859, APMCC 699, Abell 3995, Abell 2064 and RBS 171.

In addition to the clusters discussed above which are not included in the catalogue due to contamination by AGN or presence of cool-cores etc., we note that some notable nearby extended clusters are also not included in the *Planck* SZ catalogue. Indeed, the detection methods used to detect the SZ effect are not optimized for the detection of sources with scale radius θ_{500} in excess of 30 arcmin. Of the 25 clusters in this category (with $z < 0.03$) in the MCXC meta-catalogue, six are included in the *Planck* catalogue. The remaining 19 fall into the masked areas (seven out of 19, among which Perseus and Abell 1060 lie in the PS mask (Fig. 19, first two panels), and Ophiuchus and 3C 129.1 lie in the Galactic mask (Fig. 19, second two panels)) and/or have a S/N ratio below the PSZ catalogue threshold S/N = 4.5. This is the case of Virgo cluster (Fig. 19, lowest panel), which is detected in the *Planck* survey but with a signal-to-noise ratio at its position of about 3.9. Virgo's extension on the sky ($\theta_{500} = 168$ arcmin) further hampers its blind detection.

We show in Fig. 19 the reconstructed SZ signal from the MILCA algorithm (Hurier et al. 2010) for five of the "missing" extended clusters. These clusters, despite not being part of the *Planck* catalogue of SZ sources, are well detected in the *Planck* survey. They all are included in the thermal SZ map constructed from the *Planck* channel maps and presented in Planck Collaboration XXI (2013).

6.1.2. Identification with known optical clusters

A total of 182 *Planck* SZ detections are identified exclusively with optical clusters from Abell and Zwicky catalogues, and from the SDSS-based published catalogues, i.e., 26.6% of the known clusters in the *Planck* catalogue.

The *Planck* SZ candidates at S/N $\geq$ 4.5 have 111 exclusive associations with Abell or Zwicky clusters, i.e., with clusters not in any of the catalogues compiled in the MCXC meta-catalogue. In addition to these associations, 72 *Planck* detections are solely identified with clusters from the SDSS-based catalogues. These are either rich and massive systems ($R_{L^\star}$ greater than 110, $Q_{\rm SDSS} = 1$ clusters) or moderately low-richness systems ($Q_{\rm SDSS} = 2$ clusters, exhibiting hot gas as indicated by their S/N value in the RASS survey). However, not all the rich $Q_{\rm SDSS} = 1$ clusters in SDSS-based catalogues are found in the *Planck* catalogue. A total of 213 $Q_{\rm SDSS} = 1$ clusters from all four SDSS-based catalogues (201 outside the *Planck* union PS and Galactic mask) are not included in the *Planck* catalogue.

We explore why these rich clusters are not detected blindly by the SZ-finder algorithms. We first compare the richness-based masses against the X-ray luminosity-based masses of 26 of these "missing" clusters found in the MCXC meta-catalogue. We find a median ratio of 2.6 ± 1.2 for the richness-to-X-ray based masses, indicating that the richness-based masses seem to be systematically overestimated. Unlike the X-ray clusters, we thus cannot compute a reliable estimate of the expected S/N value for SZ detection of these optical clusters. We therefore directly search for the SZ signal at the positions of the 201 "missing" SDSS-clusters and found that all of them have S/N values below the *Planck* threshold, with a mean signal-to-noise of 1.6, except for three clusters. Two of these three "missing" SDSS-clusters have their S/N value from the extraction at the cluster position slightly higher than 4.5. The increase in S/N value is due to the difference in estimated background noise when centring the extraction at the cluster position as opposed to the blind detection. The third missing rich cluster is affected by contamination from CMB anisotropy, which results in a bad estimate of its size and consequently of its SZ signal.

6.1.3. Identification with known SZ clusters

The majority of the SZ clusters, from SPT or ACT, used in the validation process are low-mass systems ($M_{500}^{\rm median}$ around 2.3×10^{14} $M_\odot$). *Planck* is particularly sensitive to massive rich clusters and thus only a total of 56 of these clusters match *Planck* SZ detections, out of which 16 candidates are exclusively associated with SZ clusters[23] from ACT or SPT. Nine more ACT and SPT clusters are associated with *Planck* SZ detections between S/N = 4 and 4.5. We have searched for the SZ signal in the *Planck* data at the position of the remaining non-observed ACT/SPT clusters by extracting the SZ signal at their positions. We found that all had signal-to-noise values lower than 4.

We have also checked the redundancy of SZ detections within *Planck* by comparing the ESZ sample, constructed from 10 months of survey with a cut at Galactic latitudes of ± 14 degrees, with the present *Planck* catalogue. Of the 189 high significance ($(S/N)_{\rm ESZ} \geq 6$ ESZ detections, 184 ESZ confirmed clusters are included the present *Planck* catalogue within a distance of five arcmin from their ESZ position. The mean separation between the ESZ and present positions is of order 1.35 arcmin, within *Planck*'s positional accuracy. Their S/N values were increased by a factor 1.17 on average with respect to their $(S/N)_{\rm ESZ}$, (Fig. 20) and only four out of six of the ESZ clusters have new S/N values significantly lower than ESZ signal-to-noise threshold $(S/N)_{\rm ESZ} = 6$. They are displayed as stars in Fig. 20. Four ESZ clusters are not included the present *Planck* catalogue, they fall in, or nearby, the PS mask used for the preprocessing of the channel maps prior to running the detection algorithms. Such a mask was not utilized for the construction of ESZ sample. We choose not to a posteriori include these four "missing" ESZ clusters in the present *Planck* SZ catalogue.

6.1.4. Identification with clusters from NED or SIMBAD

As expected only a small number of clusters are identified from querying the databases, supplying identifiers for thirteen SZ *Planck* detections. This is because the information in NED and SIMBAD is redundant with that in the X-ray, optical, or SZ catalogues used for the external validation. The thirteen clusters found solely from querying the databases are found in the RASS survey but not in dedicated cluster catalogues, and thus not included in the MCXC; they are found in serendipitous *Chandra* surveys, or they are part of miscellaneous cluster catalogues.

[23] Six *Planck* clusters were confirmed from *XMM-Newton* or NTT observations and are also published in Reichardt et al. (2013).

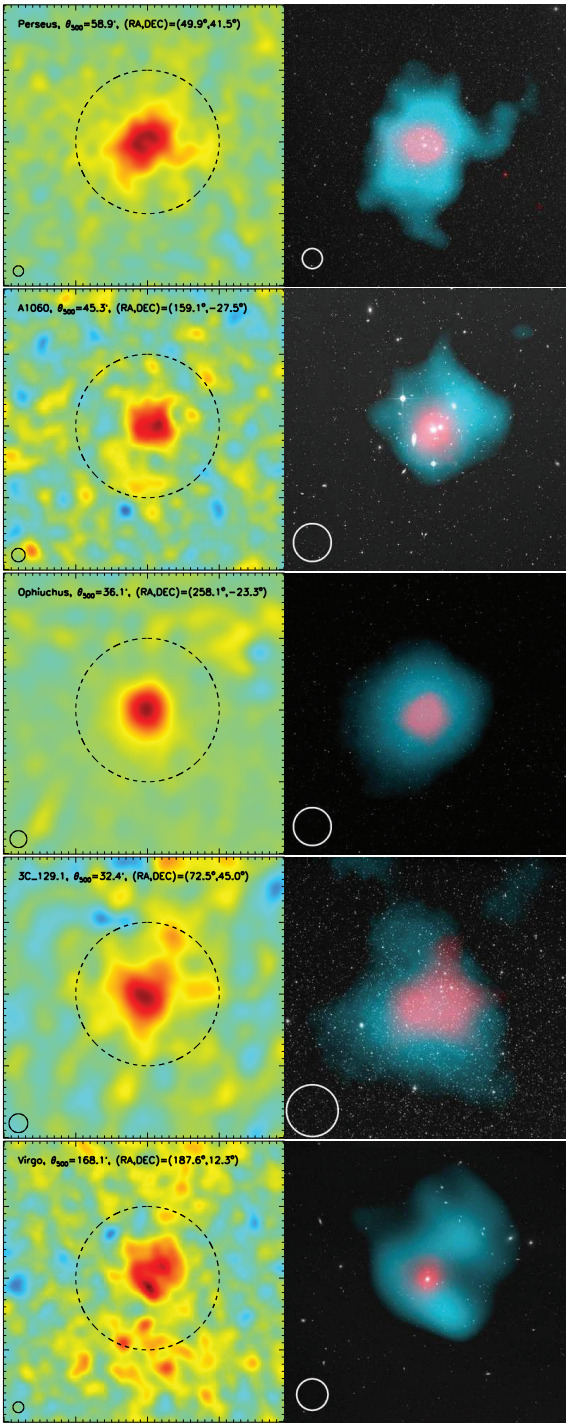


Fig. 19: Five selected nearby and extended clusters not included in the PSZ catalogue. All exhibit an extended SZ signal detected in the *Planck* survey. From top to bottom: Perseus cluster and Abell 1060 (in the point-source mask); Ophiuchus cluster and 3C 129.1 (in the Galactic mask); and Virgo cluster (below the S/N threshold of the catalogue). *Left panels:* reconstructed thermal SZ maps from the MILCA algorithm (Hurier et al. 2010). The dashed circles represent the apertures of θ_{500} from the MCXC catalogue. Each SZ-map covers an area of $4\theta_{500} \times 4\theta_{500}$. *Right panels:* composite images of the optical (DSS), X-ray (ROSAT) and SZ signal (*Planck*). The size of the composite images is adapted to optimize the display (Perseus: 2×2 square degrees; Abell 1060: 1×1 square degree; Ophiuchus: 1×1 square degree; 3C 129.1: 0.77×0.77 square degrees; Virgo: 3.84×3.84 square degrees). The solid circle in the left corner shows, both in the SZ *Planck* map and in the composite image, a 10 arcmin field of view; except for Virgo for which it shows a 30 arcmin field of view.

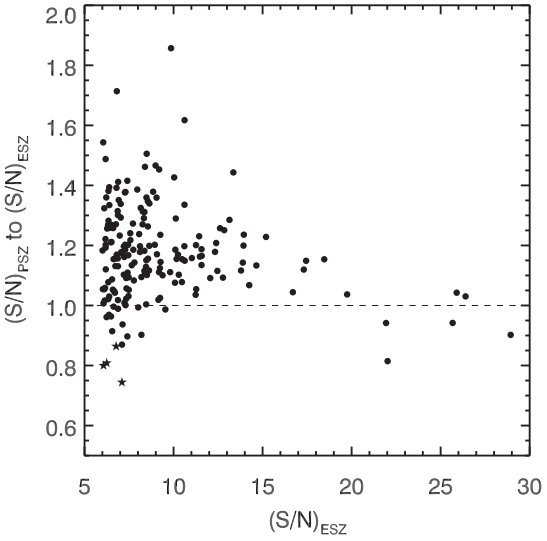


Fig. 20: Ratio of the signal-to-noise in the present *Planck* catalogue, $(S/N)_{PSZ}$, to that in the ESZ sample (Planck Collaboration VIII 2011), $(S/N)_{ESZ}$, for 184 confirmed ESZ clusters included in the *Planck* catalogue. Four clusters with signal-to-noise in the PSZ catalogue significantly smaller than the ESZ threshold ($(S/N)_{ESZ} = 6$) are shown as stars.

6.2. Newly-discovered Planck clusters and candidates

Among the 544 *Planck* SZ sources, we distinguish two categories: (1) confirmed clusters, i.e., those that have been confirmed by the follow-up programmes of the *Planck* collaboration[24] or using the SDSS galaxy catalogues. We also add eight confirmations from X-ray archival data; (2) Candidate clusters with different levels of reliability, namely, CLASS1 cluster candidates, that fulfil high-quality criteria for the SZ detection and for the associations and/or counterparts in ancillary data, CLASS2 candidate clusters, i.e., those that fulfil, on average, good-quality criteria, and CLASS3, low-reliability cluster candidates.

Confirmation from *Planck* collaboration follow-up programmes At S/N ≥ 4.5, a total of 233 *Planck* SZ detections were followed up in X-rays, optical, and SZ at the different facilities listed previously, with some observations targeted to the measurement of spectroscopic redshifts for already known clusters. In total 157 *Planck* SZ detections with S/N ≥ 4.5 were confirmed as new clusters. Some of the *Planck*-confirmed clusters were also reported in recent cluster catalogues in the optical, e.g., Wen et al. (2012) or in the SZ e.g., Reichardt et al. (2013).

The analysis of the observations of *Planck* sources by AMI yielded ten sources with strong Bayesian evidences that have clearly visible decrements and were considered as confirmed, including the confirmation of three associations with optical clusters.

For the candidates confirmed by *XMM-Newton* and by optical telescopes, redshifts from Fe lines and from photometric or spectroscopic data are available. The validation of *Planck* cluster candidates with *XMM-Newton* has shown its particular efficiency in confirming SZ candidates due both to the high sensitivity of *XMM-Newton*, allowing *Planck* clusters to be detected up to the highest redshifts (Planck Collaboration XXVI 2011), and the tight relation between X-ray and SZ properties. The detection of extended *XMM-Newton* emission and a comparison between the X-ray and SZ flux permits an unambiguous confirmation of the candidates. By contrast, confirmation in the optical may be hampered by the *Planck* positional accuracy and by the scatter between the optical observables and the SZ signal, which increase the chance of false associations. The *XMM-Newton* follow-up programme yielded 51 bona fide newly-discovered clusters, including four double systems and two triple systems. There were eight false candidates. Thirty-two of the 51 individual clusters have high-quality redshift measurements from the Fe line. The relation between the X-ray and SZ properties was used to further constrain the redshift of the other clusters; most of these redshifts were confirmed clusters using optical observations. Out of a total of 37 single clusters confirmed by *XMM-Newton*, 34 are reported in the *Planck* catalogue of SZ sources at S/N ≥ 4.5. Additionally four double systems are included in the present PSZ catalogue and were also confirmed by *XMM-Newton*.

The follow-up observations conducted with optical telescopes lead to the confirmation and to the measurement of spectroscopic or photometric redshifts (companion publications, in preparation, will present the detailed analysis and results from these follow-up). In the Northern hemisphere, 26 spectroscopic redshifts for *Planck* clusters detected at S/N ≥ 4.5 and observed at the RTT150 are reported, to date, in the PSZ catalogue. A dozen additional spectroscopic redshifts were measured for known clusters. Confirmation of 21 *Planck* SZ clusters detected above 4.5 were obtained with the ENO facilities (at INT, GTC and NOT), and robust redshift measurements were obtained for 19 of them, including 13 spectroscopic redshifts. In the Southern hemisphere, WFI observations provided photometric redshifts for 54 clusters included in the *Planck* catalogue at S/N ≥ 4.5, while 19 spectroscopic redshifts obtained with the NTT-EFOCS2 instrument are reported in the *Planck* catalogue.

Confirmation from SDSS galaxy catalogues The firm confirmation of the candidates was done through the follow-up observations for confirmation and measurement of their redshift as detailed above. However in the case of the *Planck* candidates falling in the SDSS footprint we also used the SDSS galaxy catalogues to search, as presented in Sect. 4.1.2, for galaxy overdensities associated with *Planck* SZ detections. This provides us with an estimate of the photometric redshifts, and in some cases we could retrieve spectroscopic redshifts for the BCG as well.

In this process, the major uncertainty in the associations of *Planck* SZ detections with galaxy overdensities is due to chance associations with low-richness systems or associations with diffuse concentrations of galaxies in the SDSS data. The *XMM-Newton* confirmation programmes (see Planck Collaboration Int. IV (2013) for discussion) showed that *Planck* candidates with SDSS counterparts were confirmed including PLCK G193.3−46.1 at $z \simeq 0.6$. However, the X-ray analysis of the *Planck* detections with SDSS counterparts illustrated the difficulty in distinguishing between associations of *Planck* SZ signals with massive clusters and with pre-virialized structures. In particular, in the case of extended filamentary structures or

[24] A handful of new *Planck* clusters from the ESZ sample were confirmed independently from the *Planck* collaboration by SPT (Story et al. 2011), AMI (AMI Consortium et al. 2011), Bolocam (Sayers et al. 2012a) and CARMA (Muchovej et al. 2012).

dynamically perturbed sources, an offset between the BCG position and the concentration barycentre is noted.

We considered the *Planck* SZ candidates with counterparts in the SDSS data taking into account diagnostics such as the richness/mass estimates as well as the offsets between the SZ, the BCG and the barycentre positions. We further used the outputs of the search in WISE and in RASS data, and the associated images, in order to assess the significance of the galaxy overdensity in SDSS at the position of the *Planck* candidates. For the *Planck* SZ detections where both ancillary data and SDSS barycentre/BCG positions agreed, we set that they are confirmed. We found a total of 13 such associations for which we report the photometric or the spectroscopic redshifts. It is worth noting that firm confirmation of these associations is needed and needs to be performed using either optical spectroscopic observations or X-ray observations of the *Planck* SZ detections. In the cases where the offsets between barycentre and BCG position output by the search in SDSS data were too large, and/or when other ancillary information was unable to discriminate between reliable or chance associations, we have chosen to keep the status of candidate for the *Planck* SZ detection. These cases sometimes also coincide with association of *Planck* detections with clusters from the SDSS cluster catalogues, with a quality flag $Q_{\rm SDSS} = 0$, or with confusion in the association, i.e., with positions not in agreement between counterpart and published SDSS clusters. We provide a note for all these cases in order to indicate that an overdensity in SDSS data was found.

Candidate new clusters The remaining 366 *Planck* SZ sources, not identified with previously known cluster nor confirmed by follow-up observation or ancillary data, are distributed over the whole sky (Fig. 21) and are yet to be firmly confirmed by multi-wavelength follow-up observations. They are characterized by an ensemble of quality flags defined in Sects. 4.1.1, 4.1.2, and 4.1.3 based on the systematic searches for counterparts in the public surveys during the external validation process. We further define an empirical *Planck*-internal quality flag $Q^{\rm SZ}$. It assesses the reliability of the SZ detection itself from three independent visual inspections of the nine *Planck* frequency maps, of frequency maps cleaned from Galactic emission and CMB, and of reconstructed *y*-maps or *y*-maps produced from component separation methods (e.g., Hurier et al. 2010; Remazeilles et al. 2011). Moreover, we visualize the SZ spectra from the SZ-finder algorithms and from aperture photometry measurements at the candidate positions. Finally we correlate, at the position of the *Planck* SZ candidates and within an area of 10 arcmin radius, the *y*-map to the 857 GHz channel map, as a tracer of the dust emission, and to the *Planck* mono-frequency CO map at 217 GHz (Planck Collaboration XIII 2013). The qualitative flag $Q^{\rm SZ}$ combines all this information into three values 1 to 3 from highest to lowest reliability with the following criteria:

– $Q^{\rm SZ} = 1$, i.e., high reliability: (i) Clear compact SZ source in the SZ maps; (ii) significant measurements of the SZ decrement below 217 GHz and good or reasonable detection at 353 GHz; (iii) no correlation with dust nor CO emission and no rise of the 545 and 857 GHz fluxes on the thermal SZ spectrum.
– $Q^{\rm SZ} = 2$, i.e., good reliability: (i) visible SZ detection in the SZ map or significant detection of the SZ signal below 217 GHz; (ii) contamination causing rise of the 545 GHz and possibly 857 GHz flux on the SZ spectrum without a strong correlation with dust and CO signals.
– $Q^{\rm SZ} = 3$, i.e., low reliability: (i) weak SZ signal in the *y*-maps and/or noisy SZ maps; (ii) weak or no SZ signal in the cleaned frequency maps (iii) strong correlation (≥80%) with dust and CO emission contamination with rising fluxes on the SZ spectrum at high frequencies, 353 GHz and above.

We combine the qualitative SZ quality flag with the information from the search in the all-sky surveys, RASS and WISE, for counterparts of *Planck* candidates in order to assess the overall reliability of the cluster candidates. We thus distinguish three classes of candidates:

- **CLASS1** candidates. *Highly-reliable candidates or pre-confirmed clusters:* these are the *Planck* SZ detections that have a high probability of being associated with bona fide clusters and need to fulfil high-quality criteria for SZ, RASS, and WISE detections. We retain in this category *Planck* SZ detections with high or good SZ quality flags ($Q^{\rm SZ} = 1$ or 2) and with a RASS-BSC source (not coinciding with stars) or with $(S/N)_{\rm RASS} \geq 2$, i.e., SZ detections with quality flag $Q_{\rm RASS} = 1$. The CLASS1 candidates furthermore have to fulfil a condition of high or good probability (≥80%) of being associated with an overdensity of galaxies in the WISE survey.
  We find 54 CLASS1 *Planck* candidates ranging from S/N of 4.5 to 6.3, with a median signal-to-noise ratio of 4.8. The majority of them are detected by two methods and 25.9% of them are detected only by one method. They are distributed as 26 and 28 $Q^{\rm SZ} = 1$ and 2 candidates, respectively. These candidates show significant X-ray emissions with a median $(S/N)_{\rm RASS} \simeq 3.7$ and a mean of 4.2.
- **CLASS2** candidates. *Reliable cluster candidates:* they represent 170 *Planck* SZ detections that show good or high quality criteria either in SZ or in RASS or in WISE without fulfilling all of them at once. Amongst them 61 have $Q^{\rm SZ} = 1$ and 109 have $Q^{\rm SZ} = 2$.
- **CLASS3** candidates. *Low-reliability cluster candidates:* these *Planck* SZ detections are the poor-quality, $Q^{\rm SZ} = 3$, detections. They can also be associated with good quality, $Q^{\rm SZ} = 2$, detections for which there are no good indications of the presence of an X-ray counterpart ($(S/N)_{\rm RASS} < 0.5$ and high probability of false association with FSC sources >2.5%) or a counterpart in the WISE survey (probability of association <70%).
  This class of candidates contains 142 *Planck* SZ detections with 27 and 115 SZ detection of quality $Q^{\rm SZ} = 2$ and 3, respectively.

It is worth noting that this definition of the CLASS3 *Planck* candidates is dominated by the assessment of the SZ quality complemented by information from ancillary data. In doing so we assemble in this category of candidates the SZ detections that are either false or very low quality due to contamination. Moreover, according to the statistical characterization from simulations, about 200 false detections are expected. The number of false detections could be smaller since the simulations do not reproduce the entire validation procedure, in particular omitting the cleaning from obvious false detections. Figure 22 suggests that the CLASS3 candidates are likely to be dominated by false detections. Therefore, we would like to warn against dismissing entire CLASS3 of the catalogue as populated with false detections as some CLASS3 candidates may be real clusters. For this reason, we choose not to remove these detections from the PSZ catalogue but rather flag them as low-reliability candidates. Careful

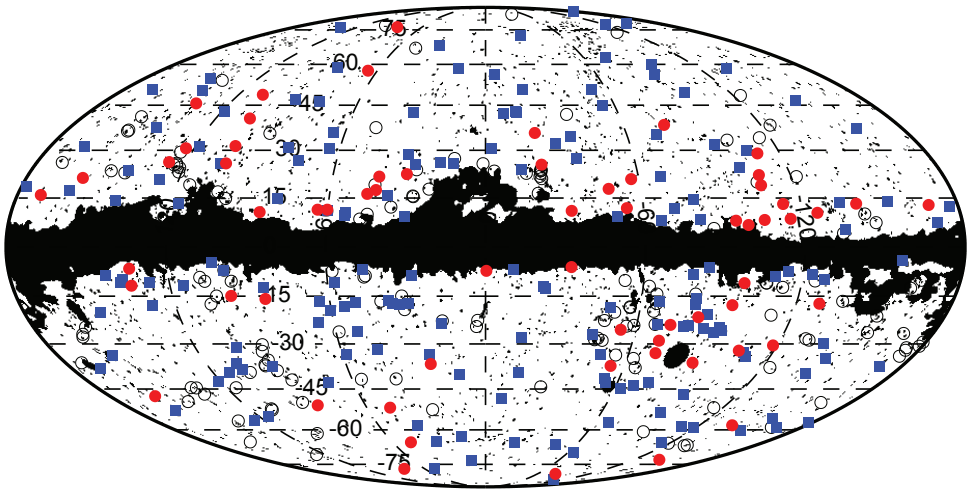


Fig. 21: Distribution of the *Planck* SZ candidates across the sky. Blue symbols represent the CLASS1 candidate clusters and red the CLASS2 candidates. The open symbols stand for the CLASS3 low-reliability SZ sources.

follow-up programmes are needed in order to separate real clusters of galaxies from false detections among the CLASS2 and CLASS3 objects.

In order to illustrate our classification defined in terms of reliability, we stack the signal in patches of 2.51 degrees across, centred at the position of the *Planck* clusters and candidates in the nine channel maps of *Planck*, removing a mean signal estimated in the outer regions where no SZ signal is expected (see Fig. 22 with the rows arranged from 30 GHz, upper row, to 857 GHz, lower row). The stacked and smoothed images are displayed for the *Planck* SZ detections identified with known clusters, CLASS1, CLASS2 and CLASS3 candidates, Fig. 22 from left to right column. We clearly see the significant detection of both the decrement and increment of the 683 *Planck* clusters and of the *Planck* candidates of CLASS1 and CLASS2. For the *Planck* SZ detections associated with bona fide clusters the increment is clearly seen at 353 and 545 GHz and is detected at 857 GHz. The smaller sample of the CLASS1 highly reliable candidates shows, in addition to the decrement at low frequency, a good detection of the increment at 353 GHz. The significance of the increment at 545 GHz is marginal and no signal is seen at 857 GHz. The case of the CLASS2 candidates (good reliability) shows that we now have lower-quality SZ detections (62% of the CLASS2 candidates have a good but not high SZ quality flag). This is illustrated by the fact that an excess emission is detected at 217 GHz, most likely due to contamination by IR sources, and both at 545 and 857 GHz where emission from dust is dominating. As for the stacked signal of the CLASS3 sample of low-reliability candidates, it does not show any significant SZ detection across frequencies, as compared to the sample of *Planck* detections identified with known clusters (Fig. 22, right column). This confirms on statistical grounds the definition of the sample dominated by definition by the low-quality SZ, $Q^{\rm SZ} = 3$, detections representing 84% of the detections in this class. Not surprisingly, the stacked signal of the CLASS3 candidates shows a large amount of contamination across all *Planck* frequencies. The low-frequency signal is dominated by radio contamination, and/or CO emission at 100 GHz, while the high-frequency signal is contaminated by emission from dust or extragalactic point sources. A more quantitative analysis is presented in Sect. 7.1.

6.3. Summary of the external validation and redshift assembly

The *Planck* catalogue of SZ sources comprises a total of 861 identified or confirmed clusters with only nine percent of them being detected by one SZ-finder algorithm. We summarize in Table 5 and Fig. 18 the results of the cluster identification. Figure 23 illustrates the status of the *Planck* SZ detections. In particular, 70.2% of the *Planck* SZ detections with S/N≥4.5 have so far been associated with clusters. The fraction increases to about 73% at S/N = 6.

We have assembled, at the date of submission, a total of 813 redshifts for the 861 identified or confirmed *Planck* clusters, which we provide together with the published *Planck* catalogue. Their distribution is shown in Fig. 24. In the process of the redshift assembly that is summarized below, especially for the already known clusters, we have favoured homogeneity for the sources of redshift rather than a cluster-by-cluster assembly of the most accurate z measure. A large fraction of the redshifts, 456 of them, shown as the dashed green histogram in Fig. 24 correspond to the spectroscopic redshifts quoted in the updated MCXC meta-catalogue (Piffaretti et al. 2011). They are associated with the *Planck* clusters identified with known X-ray clus-

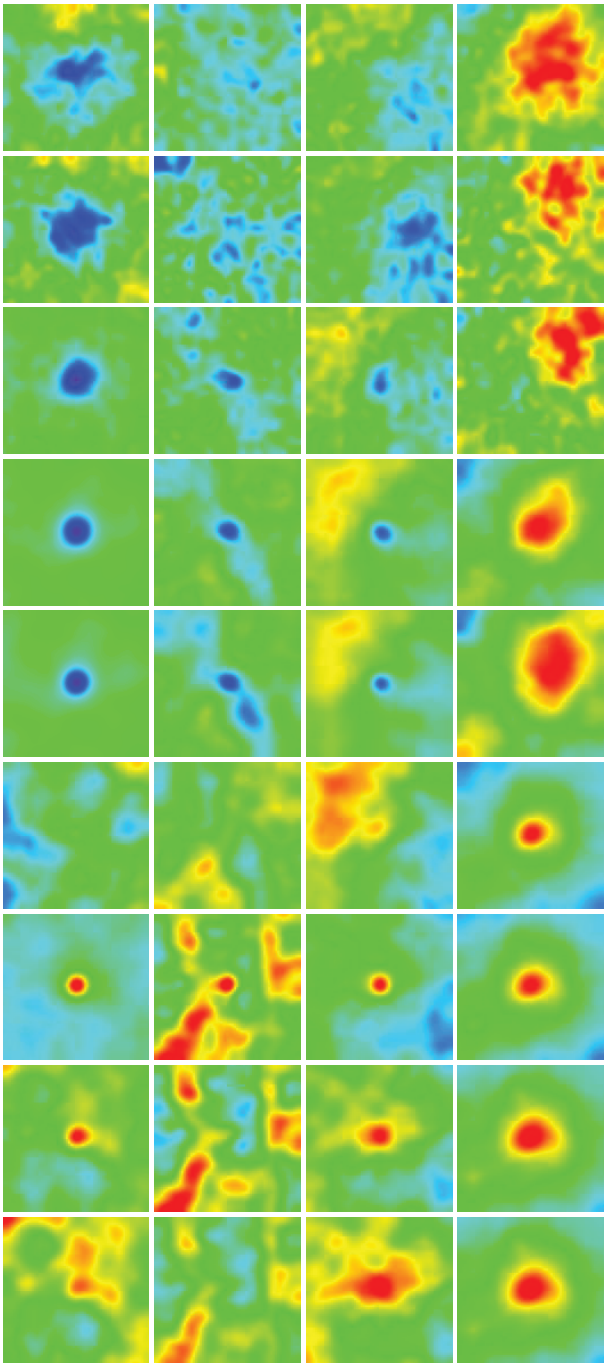


Fig. 22: Stacked signal in the nine *Planck* frequencies (30 to 857 GHz from upper to lower row). From left to right are displayed the *Planck* SZ detections identified with known clusters, the CLASS1 high-reliability *Planck* SZ candidates, the CLASS2 good-reliability *Planck* SZ candidates, and finally the CLASS3 low-reliability SZ sources. The three lowest-frequency-channel images were convolved with a 10 arcmin FWHM Gaussian kernel, whereas the remaining six highest-frequency-channel images were smoothed with a 7 arcmin FWHM Gaussian kernel.

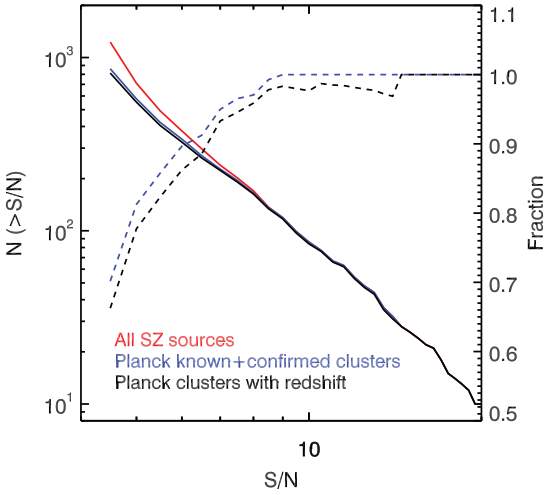


Fig. 23: Status of the *Planck* SZ sources. Left-hand-axis plots show the distribution of all *Planck* sources (in red). The blue line represents the known or new confirmed clusters and, among these, the clusters with a reported redshift measurement in black. Right-hand-axis cumulative distributions show, as a function of signal-to-noise, the fraction of known or new confirmed clusters in blue and those with a redshift in black.

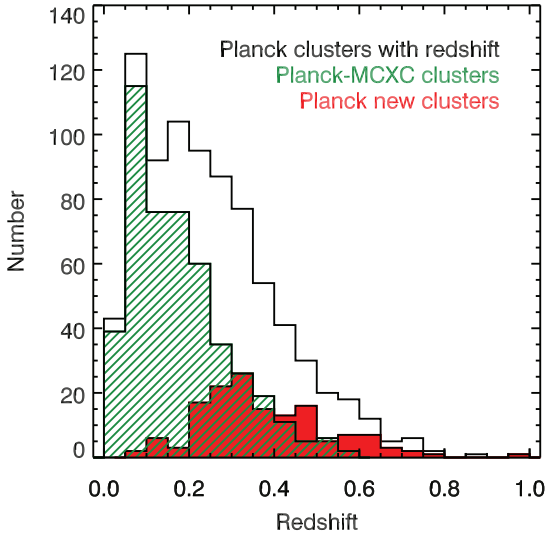


Fig. 24: Distribution of redshifts for the *Planck* SZ clusters (black line). The *Planck* clusters associated with MCXC clusters are shown in dashed green and the brand new *Planck* clusters are in the filled red histogram.

ters and they are denoted *Planck*-MCXC. For the *Planck*-MCXC clusters without reported redshifts from the MCXC, we have complemented the information with the available redshifts from NED and SIMBAD. We have further quoted when available, mainly for the MACS clusters, the estimated photometric redshifts from SDSS cluster catalogue of Wen et al. (2012). At the end only two *Planck* detections identified with MCXC clusters remain without redshifts. The redshift distribution of the *Planck* clusters identified with MCXC clusters mostly reflects that of the REFLEX/NORAS catalogues at low and moderate redshifts and the MACS clusters at higher redshifts.

For the *Planck* detections exclusively identified with Abell or Zwicky clusters, we choose to report the redshifts published in the NED and SIMBAD data bases rather than those quoted in the native catalogues. As for the *Planck* detections identified with clusters from the SDSS-based catalogues, we choose to favour homogeneity by reporting whenever possible the Wen et al. (2012) redshifts. Furthermore, we favour when available spectroscopic redshifts over photometric ones. The *Planck* detections exclusively associated with ACT or SPT clusters have published redshifts (Sifon et al. 2012; Hasselfield et al. 2013; Reichardt et al. 2013). We select in priority the spectroscopic ones when available. If not, we quote the photometric redshifts.

Finally, the follow-up observations for confirmation of *Planck* detections started in 2010 and are still ongoing. As mentioned earlier our priority was to assemble the largest possible number of confirmations and redshifts. Therefore, we did not systematically confirm the photometric redshift estimates spectroscopically. We report the obtained redshifts when available. In some cases, the new *Planck* clusters were confirmed from imaging or pre-imaging observations and the analysis is still ongoing. The spectroscopic redshifts will be updated when available. Spectroscopic redshifts for some known clusters will also be updated. A dozen *Planck* clusters were confirmed by a search in the SDSS galaxy catalogues. For these clusters, only a photometric redshift estimated by the cluster-finder algorithm of Fromenteau et al. (2013) is available and is reported.

We show in Fig. 24 the distribution of redshifts of the *Planck* clusters. The mean redshift of the sample is 0.25 and its median is 0.22. One third of the *Planck* clusters with measured redshifts lie above $z = 0.3$. The new *Planck* clusters probe higher redshifts and represent 40% of the $z \geq 0.3$ clusters. Their mean redshift is 0.38 and the median is $z = 0.35$. At even higher redshifts, $z \geq 0.5$, the *Planck* catalogue contains 65 clusters including *Planck* SZ clusters identified with WHL12's clusters (Wen et al. 2012), or with clusters from ACT and SPT, or with X-ray clusters. The *Planck* detections in this range of redshifts, 29 *Planck* new clusters, almost double the number of high redshift clusters.

The *Planck* SZ catalogue has been followed up by the *Planck* collaboration using different facilities and only a small fraction of the *Planck* candidates were observed to date. A systematic follow-up effort for the confirmation of the remaining cluster candidates will likely reveal clusters at redshifts above 0.3. As a matter of fact, very few new clusters were found below $z = 0.2$ (see Fig. 24). Such an observational programme is challenging and will most likely be undertaken by the *Planck* collaboration and by the community. It will increase further the value of the *Planck* SZ catalogue as the first all-sky SZ-selected catalogue.

7. Physical properties of *Planck* SZ clusters

The first goal of the external validation process based on the ancillary multi-wavelength data is to assess the status of the *Planck* SZ detections in terms of known clusters, brand new clusters or cluster candidates. The wealth of information assembled and used during this process also allows us to explore the properties of the *Planck* SZ clusters and candidates. We present in the following some of these properties, namely the contamination

levels of the *Planck* SZ detections, a refined measurement of the Compton Y parameter for the *Planck* clusters identified with X-ray clusters from the MCXC, an SZ-mass estimate based on a new proxy for all the *Planck* clusters with measured redshifts, and an estimate of the X-ray flux from the RASS data for the *Planck* SZ detections not included in the X-ray catalogues. This additional information associated with the *Planck* clusters and candidates derived from the validation process is summarized in the form of an ensemble of outputs given in Table C.1.

We further present an updated and extended study of the SZ versus X-ray scaling relation, confirming at higher precision the strong agreement between the SZ and X-ray measurements (within R_{500}) of the intra-cluster gas properties found by Planck Collaboration XI (2011).

7.1. Point-source contamination

Galactic and extragalactic sources, emitting in the radio or infrared domain, are known to lie in galaxy clusters and hence are a possible source of contamination for the SZ measurement (e.g., Rubiño-Martín & Sunyaev 2003; Aghanim et al. 2005; Lin et al. 2009). We address the possible contamination of the SZ flux by bright radio sources that may affect the measured signal in the direction of some of the *Planck* SZ detections. In order to do so, we searched for known radio sources in the vicinity of the *Planck* cluster candidates. In particular, we use the NVSS 1.4 GHz survey (Condon et al. 1998) and SUMMS 0.85 GHz survey (Bock et al. 1999) to identify bright radio sources within seven arcmin of the *Planck* cluster or candidate position. We assumed a spectral index $\alpha = -0.5$ for these sources to extrapolate their flux to the *Planck* frequencies. Most bright sources in NVSS and SUMSS have steeper spectral indexes (-0.6 or -0.7), so the value $\alpha = -0.5$ provides us with an upper limit in most cases. After convolving the radio sources by *Planck*'s beam, we estimate the maximum amplitude in units of μK within five arcmin of the *Planck* position. We report only those cases where this amplitude is above 5 μK in the 143 GHz channel and could thus contaminate the SZ signal. Below this value, the emission from radio sources can be considered negligible.

We find that a total of 274 *Planck* clusters and candidates, i.e., 22% of the SZ detections, are affected by such emission from bright radio sources. These clusters or candidates are identified in the PSZ catalogue and a specific note is provided. We find that the fraction of contaminated *Planck* SZ clusters identified with known X-ray, optical, or SZ clusters is also 22%. The *Planck* candidate-clusters of CLASS1 and CLASS2 are less contaminated by bright radio sources; only a fraction of 15% and 17% for CLASS1 and 2, respectively. This is due to the definition of our quality criteria for SZ detection, which results in less contamination for the high and good reliability candidates.

Another approach used to assess the contamination is based on the stacking analysis of the *Planck* clusters and candidates described in Sect. 6.2. This analysis is performed on the sample of *Planck* clusters identified with known clusters and on the sample of low-reliability CLASS3 *Planck* candidates. To do so we fit a GNFW pressure profile to the signal at 100 GHz and 143 GHz and we subtract the associated SZ signal from the stacked maps. The residual signal is then compared with a toy model for point sources ($F_\nu = S^{\rm rad}_{30}(\nu/30\,{\rm GHz})^{\alpha_{\rm rad}}$ for radio sources) and ($F_\nu = S^{\rm IR}_{857}(\nu/857\,{\rm GHz})^{\alpha_{\rm IR}}$ for IR point sources). Note that the residual signal at high frequencies is a combination of possible IR sources and IR emission from Galactic dust; the latter is not explicitly modelled in the present analysis. The PS toy models are convolved by the beam at each frequency and the signal is measured at a fixed aperture set to the FWHM of the 143 GHz channel. The average signal within this aperture is estimated for each channel before (Fig. 25, upper panel) and after (Fig. 25, lower panel) removal of the SZ signal. The black filled

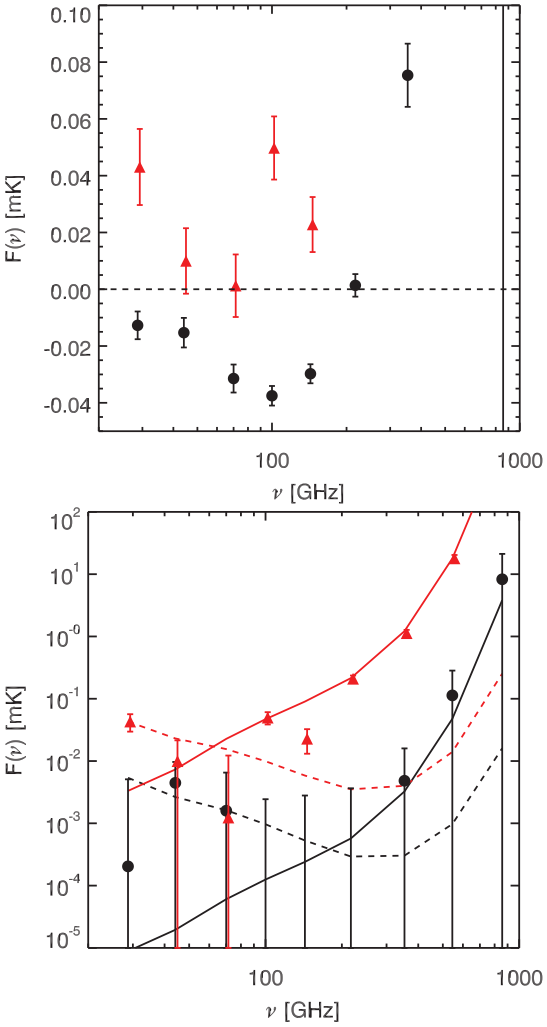


Fig. 25: Stacked spectrum for known clusters SZ fluxes across *Planck* frequency bands. Stacked fluxes are measured in an aperture equal to the FWHM of the 143 GHz channel (i.e., about 7 arcmin) for the known clusters (black filled circles) and the low-reliability CLASS3 candidates (red filled triangles). The associated uncertainties correspond to the fluctuation of the background outside the cluster region. The average signal is estimated in each channel before (upper panel) and after (lower panel) the removal of the SZ signal. The average signals expected from IR and radio sources are shown as solid and dashed lines, respectively. Red and black lines are for CLASS3 and bona fide clusters, respectively. No subtraction of an SZ signal is performed for the CLASS3 candidates.

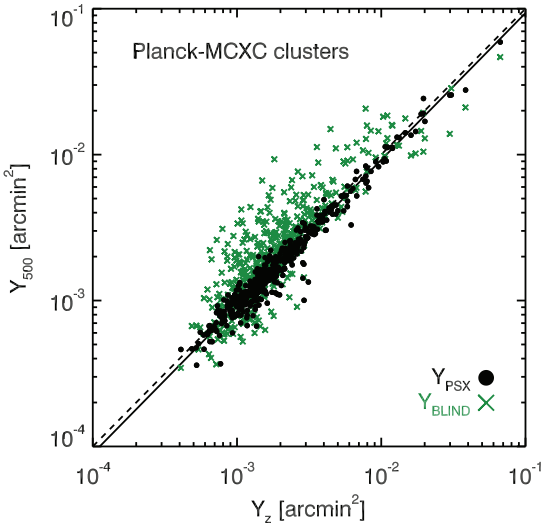


Fig. 26: Comparison of the different Y estimates for the *Planck* clusters identified with MCXC clusters. In green are the blind measured Y values and in black are the refined $Y_{500,PSX}$ measured fixing the size and positions to the X-ray values. Both are plotted as a function of the new proxy Y_z.

circles are for *Planck* SZ sources associated with known clusters and the red filled triangles stand for CLASS3 candidates. The average signal from the PS models is shown in Fig. 25 as solid (IR sources) and dashed (radio sources) lines. Red and black are for CLASS3 and bona fide clusters, respectively. The error bars correspond to the fluctuation of the background outside the cluster region. For the sample of CLASS3 candidates no SZ-signal removal was applied, since no significant detection is seen at 100 GHz or 143 GHz.

We find that the residual signal (after SZ subtraction) in the sample of known *Planck* clusters is compatible with the emission from radio sources at low frequencies with $(S_{30}^{\rm rad}, \alpha_{\rm rad}) = (14.6\,{\rm mJy}, -1)$ for the known clusters. It is also compatible with IR emission at high frequencies with a spectral index $\alpha_{\rm IR} = 2.5$, in agreement with the results of Planck Collaboration Int. VII (2013) and with $S_{857}^{\rm IR} = 0.117$ Jy. For CLASS3, where no SZ signal is subtracted, it is the full signal that is compatible with the IR emission at high frequencies, with $(S_{857}^{\rm IR}, \alpha_{\rm IR}) = (43.9\,{\rm Jy}, 2.5)$, and with radio emission from point sources with $(S_{30}^{\rm rad}, \alpha_{\rm rad}) = (117.1\,{\rm mJy}, -0.8)$.

7.2. Refined measurement of Y

While the true Y_{500} is expected to be a low-scatter mass proxy, this is not the case for the blind Y_{500}. Without a cluster-size estimate, Y_{500} cannot be accurately measured. Moreover, the blind SZ flux is biased high on average, because the size is overestimated on average. This effect is amplified by the non-linear nature of the size–flux degeneracy, with a larger effect of size over-estimation than size under-estimation. This behaviour, first identified and discussed in Planck Collaboration VIII (2011) and Planck and AMI Collaborations (2013), hampers the direct use of the blind SZ fluxes as a mass proxy. As shown in Planck Collaboration VIII (2011), this degeneracy calls for a refined measurement of the SZ signal. In this section, we present two ways of refining the Y measurement. Both are based on fixing the cluster size in two cases, by setting it equal to the X-ray estimated size or by using the redshift information when available. The outputs of the refined measurement are provided as additional information complementary to the catalogue of *Planck* SZ detections (see Appendix C and Table C.1).

7.2.1. Y at fixed X-ray size and position

As shown by Planck Collaboration VIII (2011), the size–flux degeneracy can be broken by introducing a higher-quality estimate of the cluster size θ_{500}. This prior is directly provided by X-ray observations using an X-ray mass proxy such as $Y_{\rm X}$ or the luminosity $L_{\rm X}$. Resorting to estimates of the cluster size from optical richness is also possible, but suffers from the large scatter in richness–mass relation, as discussed previously.

A detailed investigation of the effects of fixing the cluster size was presented in Planck Collaboration XI (2011, Appendix A). Following this approach, and for the *Planck* detections identified with clusters from the MCXC meta-catalogue, we have adopted the R_{500} and z values reported in Piffaretti et al. (2011) as priors to re-extract at the X-ray position the SZ signal denoted $Y_{500,PSX}$ assuming the Arnaud et al. (2010) pressure profile (see Table C.1). The comparison between the blind Y_{500} and refined $Y_{500,PSX}$ (Fig. 26) shows that both the scatter and the offset are significantly reduced by the refined SZ measure. The SZ re-extraction at X-ray position and fixing the size to the X-ray derived size provides an unbiased estimate of the SZ signal. However, as stressed in Planck Collaboration XI (2011, Appendix A), the MCXC cluster size derivation involves the M_{500}–$L_{\rm X,500}$ relation, which exhibits a non-negligible scatter. This leads to a remaining systematic discrepancy between the expected Y value from X-ray measurements and the actual SZ flux derived from the *Planck* data. The use of the $Y_{\rm X}$ proxy does not suffer from such an effect, but high-quality X-ray data permitting the use of such a quantity are not available for a large number of clusters (see Sect. 7.5 for the presentation of a sample of *Planck* SZ clusters with high-quality X-ray data).

7.2.2. Y from the $Y(\theta)$–M relation

The size–flux degeneracy can further be broken, as proposed by Arnaud et al. (2013), using the M_{500}–$D_{\rm A}^2 Y_{500}$ relation itself that relates θ_{500} and Y_{500}, when z is known. Then Y_{500} is derived from the intersection of the M_{500}–$D_{\rm A}^2 Y_{500}$ relation and the size–flux degeneracy curve. A detailed description of the method and the comparison of results in terms of bias and scatter can be found in Arnaud et al. (2013).

The derived Y_{500} parameter is denoted Y_z (since it involves a measurement of the Compton Y signal for clusters with measured redshift z). It is the SZ mass proxy Y_z that is equivalent to the X-ray mass proxy $Y_{\rm X}$. Y_z is computed for all the 813 *Planck* clusters with measured redshifts. We use Malmquist-bias-corrected scaling relation between mass and Y given in Planck Collaboration XX (2013)

$$E^{-2/3}(z)\left[\frac{D_{\rm A}^2(z)\,Y_{500}}{10^{-4}\,{\rm Mpc}^2}\right] = 10^{-0.19}\left[\frac{M_{500}}{6\times10^{14}\,{\rm M}_\odot}\right]^{1.79}, \quad (5)$$

with $E^2(z) = \Omega_{\rm m}(1+z)^3 + \Omega_\Lambda$ computed in the fiducial ΛCDM cosmology.

In Fig. 26, the refined Y_{500} value, measured fixing the size and position to the X-ray values $Y_{500,PSX}$, is compared to the

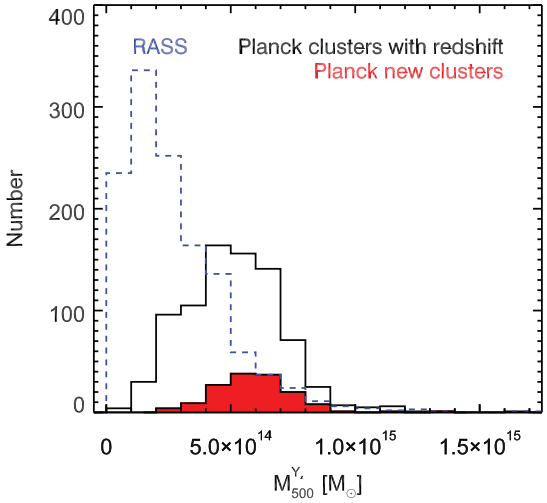


Fig. 27: Distribution of masses for the *Planck* SZ clusters, known or new confirmed clusters (solid black line), compared to the distribution of masses from the RASS-based cluster catalogues (dashed blue line). The masses for the MCXC clusters are estimated from the luminosity–mass relation. The masses for the *Planck* clusters are computed using the SZ-proxy. The filled red histogram shows the distribution of the newly-discovered *Planck* clusters.

blind Y as a function of the derived Y_z proxy. We see that the scatter and the offset are significantly reduced.

Under the two hypotheses of cosmology and scaling relation, Y_z provides the best estimate of Y_{500} for the *Planck* SZ clusters and conversely a homogeneously-defined estimate of an SZ-mass, X-ray calibrated, denoted $M_{500}^{Y_z}$. For the ensemble of *Planck* clusters with measured redshifts, the largest such sample of SZ-selected clusters, we show in Fig. 27 the distribution (black solid line) of the masses obtained from the SZ-based mass proxy. The distribution of the SZ masses is compared with those of the RASS clusters (dashed blue line) computed from the X-ray luminosity $L_{X,500}$. The mean and median masses of the *Planck* clusters are 3.3 and $3.5 \times 10^{14}\,M_\odot$, respectively. The *Planck* SZ catalogues contains all the massive clusters of the RASS catalogues. Interestingly, the distribution of newly-discovered *Planck* clusters extends to higher masses with a median mass of $5.7 \times 10^{14}\,M_\odot$. Besides providing a homogeneous estimate of the masses from an SZ proxy for the largest SZ selected sample of clusters, we show that *Planck* detections significantly extend the mass range in the high-mass region up to $1.6 \times 10^{15}\,M_\odot$.

7.3. *M–z* distribution and comparison with other surveys

Based on the masses derived from the SZ-proxy, we illustrate for MMF3 the *M–z* distribution of *Planck* SZ clusters detected over 83.7% of the sky. We show in all panels of Fig. 28 the limiting mass $M_{\rm lim}$ computed following Planck Collaboration XX (2013) for three values of the completeness: 20% (solid line); 50% (dashed line); and 80% (dotted line). The upper left panel exhibits the *Planck* clusters, with redshifts, detected by MMF3 at S/N≥4.5. The mass limit corresponds to the average limit computed from the noise over the 83.7% sky fraction used by the SZ-finder algorithm. The resulting $M_{\rm lim}$ is not representative of the inhomogeneity of the noise across the sky (see Fig. 3). We therefore show the limiting mass in three areas of the sky (Fig. 3): the deep-survey area (upper right panel); the medium-deep survey area (lower left panel); and the shallow-survey area (lower right panel). The lines indicate the limit at which clusters have C% chances to be detected (C being the completeness value). We clearly see that whereas the average $M_{\rm lim}$ at 20% completeness does not fully represent the SZ detections by MMF3, the limiting masses in different survey depths are more representative of the detection process. We further note that except at low redshifts, $z < 0.3 - 0.4$, the *Planck* cluster distribution exhibits a nearly redshift-independent mass limit with a cut that varies according to the survey depth.

It is worth examining the distribution of the *Planck* SZ clusters in the *M–z* plane and comparing it to that of other catalogues. For illustration, we compare to an X-ray selected sample, namely REFLEX-I, on the one hand (Fig. 29, right panel green open circles) and to the large-area SZ-selected cluster catalogues by ACT (Hasselfield et al. 2013) and SPT (Reichardt et al. 2013), on the other hand (Fig. 29, red open symbols). In this comparison we report, for the ACT clusters (open squares), the so-called UPP (Universal Pressure Profile) masses given in Hasselfield et al. (2013).

The range of redshifts covered by the *Planck* SZ sample, from $z = 0.01$ to about 1 with 67% of the clusters lying below $z = 0.3$, is quite complementary to the high redshift range explored by ACT and in by SPT. For the comparison of the mass distribution we take advantage of our newly-proposed SZ-mass estimate, derived from Y_z, which provides us with a homogeneous definition of the masses over the whole range of *Planck* SZ clusters with measured redshifts. The *Planck* clusters populate the full redshift range and they quite nicely fill a unique space of massive, $M \geq 5 \times 10^{14}\,M_\odot$, and high redshift $z \geq 0.5$ clusters, as shown in Fig. 29. This contrasts with the SZ clusters detected in 720 square degrees of SPT observations and those of ACT observations, which are dominated, as shown in Fig. 29 left panel, by lower-mass higher-redshift clusters (up to $z \sim 1.3$). The combination of *Planck* and SPT/ACT catalogues samples the *M–z* space in a complementary manner. Clearly the all-sky nature of the *Planck* makes the most massive clusters preferentially accessible to *Planck* whereas the highest redshift clusters, $z \geq 1$, are accessible to SPT.

Very few massive high-redshift clusters exist in the X-ray catalogues, as seen in Fig. 29 (right panel open blue squares). The all-sky NORAS/REFLEX catalogues (Böhringer et al. 2000, 2004) are limited to $z = 0.45$, a result of the $(1+z)^4$ surface brightness dependence of the X-ray detection limit (Fig. 29, right panel solid green line). The smaller-area MACS sample, based on systematic follow-up of *ROSAT* bright sources (Ebeling et al. 2007), contains a dozen clusters at $z \geq 0.5$. The 400SD sample (Burenin et al. 2007), based on serendipitous detections in 400 deg^2 of ROSAT pointed observations, contains only two clusters with $M \geq 5 \times 10^{14}\,M_\odot$ and $z \geq 0.5$. Finally, only a couple of clusters in the range $M \geq 5 \times 10^{14}\,M_\odot$ are found in the *XMM-Newton* based serendipitous cluster samples (XCS, Mehrtens et al. (2012); XMM-LSS, Pacaud et al. (2007); XDCP, Fassbender et al. (2011)). By contrast to an X-ray selected cluster catalogue, the *Planck* detection-limit, illustrated for the medium-deep survey zone and shown in Fig. 29 (right panel solid black line), has a much shallower dependence on redshift and is quasi-redshift independent above $z = 0.4$. The difference in cluster selection starts at redshifts $z \geq 0.2$. As a re-

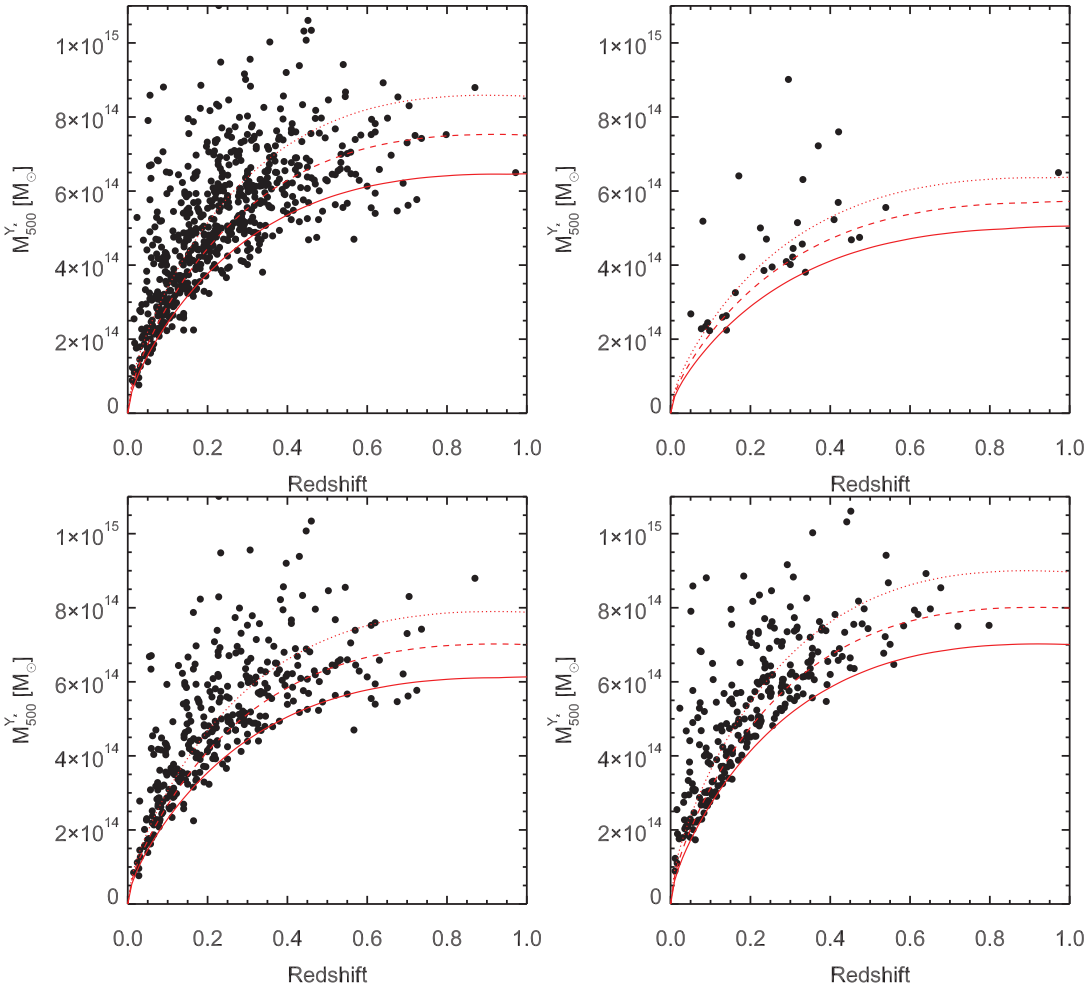


Fig. 28: Mass limit illustrated for SZ detections by MMF3 algorithm. *Upper left:* average mass limit computed from the average noise over the sky. *Upper right:* same for the deep survey zone corresponding to 2.7% sky coverage centred at the Ecliptic polar regions. *Lower left:* same for the medium-deep survey area covering 41.3% of the sky. *Lower right:* same for the shallow-survey area covering 56% of the sky. In each panel, only detections in the corresponding areas are plotted. The lines dotted, dashed and solid lines show the *Planck* mass limit at 80, 50 and 20% completeness, respectively.

sult of the quasi-redshift independent mass-selection of SZ surveys, *Planck* probes deeper than the X-ray selection. This is also seen in the overall distribution of redshifts of the *Planck* clusters, Fig. 24.

This leaves the *Planck* SZ catalogue as the deepest all-sky catalogue spanning the broadest cluster mass range from 0.1 to 1.6×10^{15} $M_\odot$, and particularly adapted to the detection of rare very massive clusters in the tail of the distribution in the range $M \geq 5 \times 10^{14}$ $M_\odot$ and $z \geq 0.5$.

7.4. X-ray flux of the Planck clusters and candidates

For all *Planck* SZ detections, we estimated the unabsorbed fluxes at Earth in the [0.1–2.4] keV band (as in the MCXC) measured in an aperture of five arcmin. The aperture is centred on the *Planck* candidate position, except for candidates associated with a BSC source, for which we adopt the X-ray position, since the BSC source is very likely the counterpart (Planck Collaboration Int. IV 2013). The conversion between the RASS count rate in the hard band and flux is performed using an absorbed thermal emission model with the $N_{\rm H}$ value fixed to the 21 cm value. The conversion depends weakly on temperature and redshift and we assumed typical values of $kT = 6$ keV and $z = 0.5$. Planck Collaboration Int. IV (2013) compared such flux estimates with precise *XMM-Newton* fluxes measured within R_{500}, S_{500}, for candidates confirmed with the *XMM-Newton* follow-up programme. These clusters lie in the range $0.1 < z < 0.9$ and the $0.3 \times 10^{-12} < S_{500} < 6 \times 10^{-12}$ erg s^{-1} cm^{-2} flux range. The

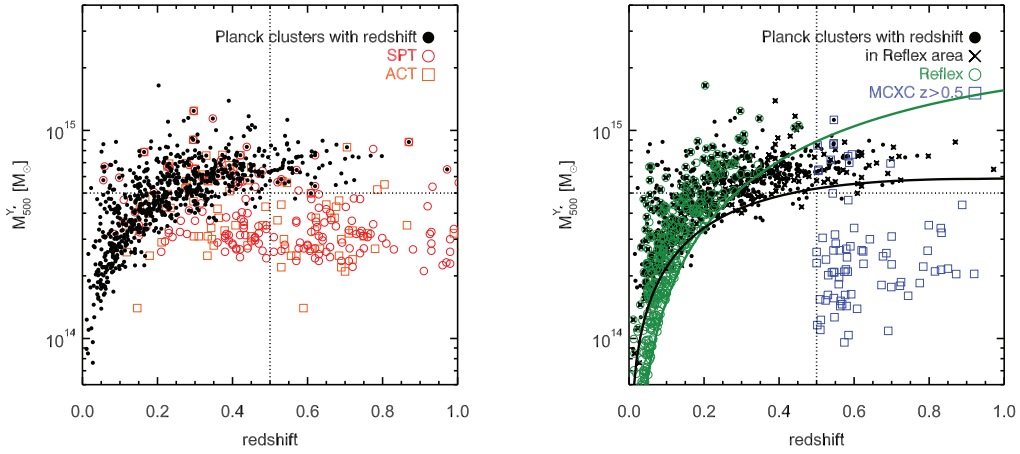


Fig. 29: *Left panel:* distribution in the *M*–*z* plane of the *Planck* clusters (filled circles) compared with the SPT clusters (open light red circles) from Reichardt et al. (2013) and ACT catalogue (open red squares) from Hasselfield et al. (2013). *Right panel:* distribution in the *M*–*z* plane of the *Planck* clusters (black symbols) as compared to the clusters from the REFLEX catalogue (green open circles) Böhringer et al. (2004). The black crosses indicate the *Planck* clusters in the REFLEX area. The open blue squares represent clusters from the MCXC catalogue with redshifts above $z = 0.5$. The green solid line shows the REFLEX detection limit whereas the black solid line shows the *Planck* mass limit for the medium-deep survey zone at 20% completeness.

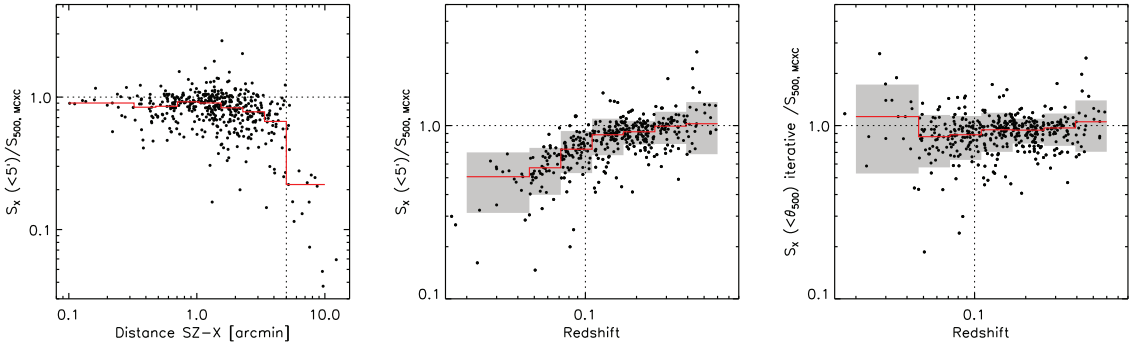


Fig. 30: Ratio between RASS flux, computed in an aperture of five arcmin in radius centred on the *Planck* position, and MCXC value for *Planck* candidates identified with MCXC clusters. The fluxes are computed in the [0.1–2.4] keV band at Earth and corrected for absorption. S_{500} is the flux corresponding to the luminosity within R_{500} published in the MCXC catalogue. *Left panel:* the ratio is plotted as a function of distance between the *Planck* and X-ray positions; *Middle panel:* same, as a function of cluster redshift, for distances smaller than five arcmin; *Right panel:* same as middle panel, for RASS flux within R_{500} derived from the aperture flux, using the MCXC iterative procedure based on the L_{500}–M_{500} relation and the REXCESS gas density profile (Piffaretti et al. 2011). The red line is the median ratio in distance or redshift bins with the grey area corresponding to $\pm 1\,\sigma$ standard deviation in each bin.

RASS aperture fluxes were found to underestimate the "true" flux by about 30 %.

Figure 30 extends this comparison further to all the *Planck* SZ detections identified with MCXC clusters. Piffaretti et al. (2011) published homogenized L_{500} and R_{500} values derived from the flux given in the original catalogues in various apertures, using an iterative procedure based on the REXCESS L_{500}–M_{500} relation and gas density profile shape. We simply computed S_{500} from L_{500}, taking into account the K-correction at the cluster redshift, but neglecting its variation with temperature.

Although derived from *ROSAT* survey data as our present flux estimate, S_{500} values from the MCXC are expected to be more accurate due to: (i) optimum choice of the X-ray centre; (ii) higher S/N detection; (iii) more sophisticated flux extraction adapted to data quality and source extent (e.g., growth curve analysis); and (iv) use of R_{500} rather than a fixed aperture. Not surprisingly, the ratio between the present flux estimate and the MCXC value decreases with increasing offset between the *Planck* position and X-ray position (Fig. 30, left panel). The ratio drops dramatically when the distance is larger than five arcmin, i.e., when the X-ray peak lies outside the integration aperture.

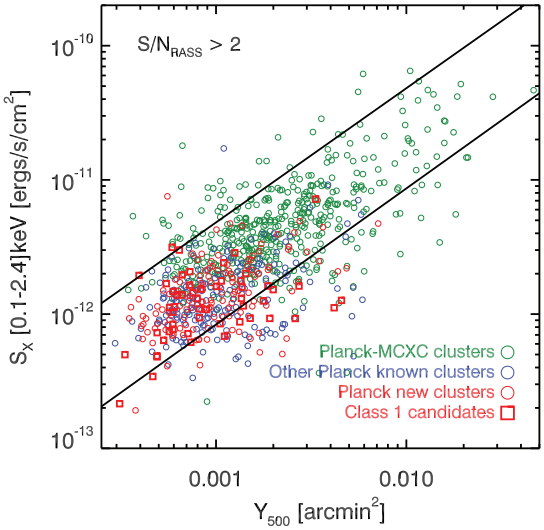


Fig. 31: X-ray unabsorbed flux versus SZ flux. For *Planck* SZ detections identified with MCXC clusters (open green circles), the X-ray flux is estimated from L_{500}. For other *Planck* SZ detections, the flux is derived from RASS count-rate in a five-arcmin aperture (see Sect. 4.1.1). *Planck* new clusters and CLASS1 candidates are shown as open red circles and squares, respectively. The two lines corresponds to the expected L_{500}–Y_{500} relation (Arnaud et al. 2010) at $z = 0.01$ and $z = 1$, respectively.

Those are rare cases, 18 nearby clusters ($z < 0.1$ with a median value of $z = 0.05$), for which a physical offset likely contributes to the overall offset. When these cases are excluded, the median ratio is 0.85 and depends on redshift (Fig. 30, middle panel); it significantly decreases with decreasing redshift below z of 0.1. The median ratio is 0.65 and 0.92, with a standard deviation of 0.10 and 0.15 dex, below and above $z = 0.1$, respectively. This is mostly due to the choice of a fixed aperture that becomes too small as compared to R_{500} at low z. If we apply the same iterative procedure used by Piffaretti et al. (2011) to estimate S_{500} from the aperture flux, the resulting value is consistent on average with the MCXC value at all redshifts (Fig. 30, right panel). The dispersion is slightly increased. The aperture unabsorbed fluxes are thus reliable estimates of the X-ray fluxes above $z > 0.1$ on average.

Figure 31 shows the X-ray flux as function of Y_{500} for *Planck* candidates identified with known clusters, for the confirmed new *Planck* clusters and for the CLASS1 candidates. For *Planck* detections identified with MCXC clusters we plot the more precise published S_{500} value. All three categories of sources behave in a similar manner in good agreement with the range of redshifts probed by the sample. In this respect CLASS1 candidates do not exhibit any departure with respect to the known or confirmed clusters. We provide the X-ray fluxes for the *Planck* clusters and candidates that are not identified with MCXC clusters (see Appendix C and Table C.1). For the *Planck* cluster with MCXC identifier, we refer the reader to the RASS catalogue outputs or to the homogenized MCXC meta-catalogue. The main limitation of the aperture unabsorbed fluxes is the statistical precision on the RASS estimate (most of the *Planck* SZ detections not identified with MCXC clusters have low $(S/N)_{RASS}$ values) and the relatively large scatter (±30% standard deviation). For $z < 0.1$ clusters, and if the RASS detection is reasonably good a more precise procedure is recommended, such as an adapted growth curve analysis, on a case-by-case basis.

7.5. Scaling relations between SZ and X-ray quantities

A fundamental scaling relation is that between Y_{500} and its X-ray analogue, Y_X. Introduced by Kravtsov et al. (2006), Y_X is the product of $M_{g,500}$, the gas mass within R_{500}, and T_X, the spectroscopic temperature outside the core [25]. From the fact that the gas density profile used to compute $M_{g,500}$ is derived from deprojection of the X-ray surface brightness profile, and that the X-ray emission depends on the square of the density, the ratio of these two quantities is

$$\frac{D_A^2\, Y_{500}}{C_{XSZ}\, Y_X} = \frac{1}{Q}\frac{\langle n_e T\rangle_{R_{500}}}{\langle n_e\rangle_{R_{500}} T_X} \qquad (6)$$

$$Q = \frac{\sqrt{\langle n_e^2\rangle_{dr}}}{\langle n_e\rangle_{dr}},$$

where the angle brackets denote volume-averaged quantities, and Q is the clumpiness factor at the scale of the radial bins used to derive the density profile. The numerical constant $C_{XSZ} = \sigma_T/(m_e c^2\, \mu_e\, m_p) = 1.416 \times 10^{-19}\,\mathrm{Mpc^2\,(M_\odot\,keV)^{-1}}$. The ratio thus depends only on the internal structure of the intra-cluster medium.

The properties of the Y_X–Y_{500} relation, in particular its variation with mass and redshift and the dispersion about the mean relation, are important probes of the physics of cluster formation.

7.5.1. Data set

Here we extend the study of a sample of 62 clusters from the *Planck*–ESZ sample with good quality *XMM-Newton* archive data presented in Planck Collaboration XI (2011, hereafter PEPXI). This study found $D_A^2\, Y_{500}/C_{XSZ}\, Y_X = 0.95 \pm 0.03$, in a good agreement with REXCESS prediction, 0.924 ± 0.004, of Arnaud et al. (2010).

All 62 objects in the PEPXI sample are included in the present catalogue. We further add 40 clusters from the catalogue, including nine additional objects from the *XMM-Newton* archival study of *Planck*-detected LoCuSS systems presented by Planck Collaboration Int. III (2013), and the 31 *Planck*-discovered clusters with good redshift estimates ($Q_z = 2$) confirmed with the *XMM-Newton* (Planck Collaboration IX 2011; Planck Collaboration Int. I 2012; Planck Collaboration Int. IV 2013). The total sample thus consists of 102 clusters.

For each object, Y_X and the corresponding R_{500} value were estimated simultaneously by iteration about the M_{500}–Y_X relation of Arnaud et al. (2010),

$$E^{2/5}(z) M_{500} = 10^{14.567}\left[\frac{Y_X}{2\times 10^{14}\,\mathrm{M_\odot\,keV}}\right]^{0.561} \mathrm{M_\odot}. \qquad (7)$$

In the present study, we focus on the physical relation between Y_{500} and Y_X. While these quantities must be estimated within the same radii, the exact value of R_{500} is irrelevant as the radial dependence of the Y_{500}/Y_X ratio is negligible. We thus propagated only the measurement uncertainties on the temperature and gas

[25] Here we use the temperature measured in the [0.15–0.75] R_{500} aperture.

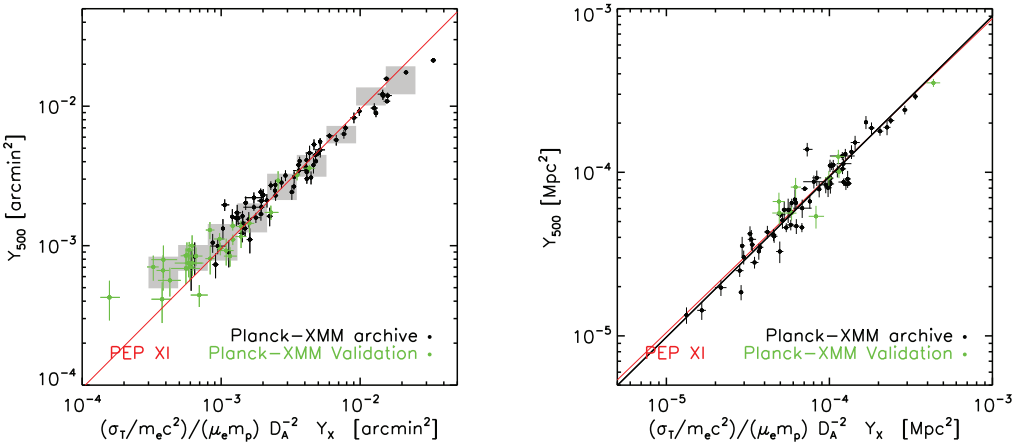


Fig. 32: Relation between the Comptonization parameters Y_{500}, and the normalized Y_X parameter for a sub-sample of the present catalogue. Black points show clusters in the *Planck*-ESZ sample with *XMM-Newton* archival data presented by Planck Collaboration XI (2011) and additional LoCuSS clusters studied by Planck Collaboration Int. III (2013). Green points represent new *Planck* clusters confirmed with *XMM-Newton* (Planck Collaboration IX 2011; Planck Collaboration Int. I 2012; Planck Collaboration Int. IV 2013). The red line denotes the scaling relations of Planck Collaboration XI (2011). *Left panel:* relation in units of arcmin^2 where Y_{500} is extracted using the Arnaud et al. (2010) pressure profile. The grey area corresponds to median Y_{500} values in Y_X bins with $\pm\sigma$ standard deviation. *Right panel:* scaling relation between the intrinsic Compton parameter, $D_A^2 Y_{500}$, and Y_X for the sub-sample of S/N > 7 clusters used in the cosmological analysis. The data are corrected for Malmquist bias, and Y_{500} is extracted using the *Planck* pressure profile (see text). The black line is the best-fit power-law relation.

Table 6: The Y_{500}–Y_X relation. Column(1): sample under consideration; Column(2): Malmquist bias correction; Column (3): Pressure profile shape used in Y_{500} extraction; Column (4): number of clusters in the sample; Column (5-6) Slope and normalization of the best-fit relation expressed as $Y_{500}/Y_p = A\,(Y_X/Y_p)^\alpha$, using BCES orthogonal regression. The pivot is $Y_p = 10^{-4}$Mpc^2; Column (7-8) Intrinsic and raw scatter around the best-fit relation; Column (9): mean ratio in logarithm, $\Delta(\log(Q))$=log(Y_{500}/Y_X); Column (10-11): corresponding intrinsic scatter and raw scatter. The scatters are error-weighted values. The best estimate is indicated in bold face. The last line gives the REXCESS prediction (Arnaud et al. 2010).

Data				Power-law Fit				Mean ratio		
Sample	MB	P Profile	N_c	$A \times 10^2$	α	$\sigma_{\log}^{\rm int} \times 10^2$	$\sigma_{\log}^{\rm raw} \times 10^2$	$\Delta\,log(Q)$	$\sigma_{\rm int} \times 10^2$	$\sigma_{\rm raw} \times 10^2$
PEPXI	N	A10	62	-2.0 ± 1.0	0.960 ± 0.040	10.0 ± 1.0	-	-0.022 ± 0.014	...	...
ESZ	N	A10	62	-2.2 ± 1.1	0.966 ± 0.034	7.2 ± 1.1	8.2 ± 1.0	-0.023 ± 0.011	7.3 ± 1.1	8.5
ESZ	Y	A10	62	-3.0 ± 1.1	0.975 ± 0.035	7.1 ± 1.1	8.2 ± 1.0	-0.031 ± 0.011	7.2 ± 1.1	8.4
S/N > 7	Y	A10	78	-2.4 ± 1.0	0.972 ± 0.029	6.9 ± 1.1	8.1 ± 0.9	-0.024 ± 0.010	6.9 ± 1.0	8.3
Cosmo	Y	A10	71	-1.9 ± 1.1	0.990 ± 0.032	7.2 ± 1.2	8.3 ± 1.0	-0.021 ± 0.010	6.9 ± 1.0	8.3
Cosmo	Y	A10+err	71	-1.9 ± 1.1	0.987 ± 0.031	6.3 ± 1.1	7.9 ± 0.9	-0.019 ± 0.010	6.5 ± 1.1	8.2
Cosmo	**Y**	**PIP-V**	**71**	$\mathbf{-2.6 \pm 1.0}$	$\mathbf{0.981 \pm 0.027}$	$\mathbf{6.6 \pm 1.2}$	$\mathbf{7.8 \pm 1.0}$	$\mathbf{-0.027 \pm 0.010}$	$\mathbf{6.6 \pm 1.0}$	**8.0**
REXCESS	X-ray prediction							-0.034 ± 0.002		

mass profiles, fixing the aperture to R_{500}. We ignored the statistical and systematic uncertainties on the M_{500}–Y_{500} relation itself.[26] Similarly Y_{500} was re-extracted at the X-ray position with size fixed to X-ray size. Its uncertainty corresponds to the statistical error on the SZ signal. The results are summarized Table 6, with the best estimate indicated in bold face.

[26] These must however be taken into account when using Y_{500} or Y_X as a mass proxy, e.g., when calibrating the Y_{500}–M_{500} relation from combining the M_{500}–Y_X relation and the relation between Y_{500} and Y_X (or equivalently M_{500}). This calibration is extensively addressed in the Planck Collaboration XX (2013).

7.5.2. The best-fit Y_{500}–Y_X relation

The Y_{500}–Y_X scaling relation for the full sample is shown in units of arcmin^2 in Fig. 32. At high flux the points follow the PEP XI relation. The slope and normalization are determined at slightly higher precision, due to the better quality SZ data. The derived intrinsic scatter (Table 6) is significantly smaller, a consequence of the propagation of gas mass profile errors in the Y_X error budget, which was neglected in our earlier study.

The relation levels off at around $Y_X = 5 \times 10^{-4}$ arcmin^2, with a bin average deviation increasing with decreasing Y_X (Fig. 32 left panel). This is an indication of Malmquist bias, as noted by Planck Collaboration Int. I (2012). Full correction of this bias when fitting scaling relations involves drawing mock catalogues according to the cluster mass function, to which the sample selection criteria are then applied. The present sample is a small

subset of the full S/N ≥ 4.5 *Planck* catalogue and thus such a procedure cannot be applied. To minimize bias effects we will only consider high S/N detections, S/N > 7. To correct for the residual bias, we adapted the approach proposed by Vikhlinin et al. (2009). Before fitting the Y_{500}–Y_X relation, each individual Y value was divided by the mean bias, b, given by

$$\ln b = \frac{\exp\left(-x^2/2\sigma^2\right)}{\sqrt{\pi/2}\,\mathrm{erfc}\left(x/\sqrt{2}\sigma\right)}\,\sigma\,, \quad (8)$$

where $x = -\log(Y/Y_{\rm min})$, $Y_{\rm min}$ being the flux threshold corresponding to the signal-to-noise cut, $(\mathrm{S/N})_{\rm cut}$. At the location of the cluster, $Y/Y_{\rm min} = (\mathrm{S/N})/(\mathrm{S/N})_{\rm cut}$. Here σ is the log-normal dispersion at fixed Y_X. We took into account both the intrinsic dispersion $\sigma_{\rm int}$, estimated iteratively, and the statistical dispersion, given by $\sigma = \sqrt{\ln\left[((\mathrm{S/N})+1)/(\mathrm{S/N})\right]^2 + [\ln 10\,\sigma_{\rm int}]^2}$. The correction decreases the effective Y_{500} values at a given Y_X, an effect that is larger for clusters closer to the S/N threshold; i.e., low-flux objects. The net effect on the scaling relation is small, giving a $0.7\,\sigma$ decrease of the normalization and a slight steepening of the power-law slope (Table 6).

The slope and normalization of the relation are robust to the inclusion of newly-discovered *Planck* clusters. The results derived from the extended sample of 78 clusters with S/N > 7 agree with those obtained for the updated XMM-ESZ sample within $0.5\,\sigma$ (Table 6). They are also in agreement with the subsample of 71 S/N > 7 clusters included in the cosmological sample discussed by Planck Collaboration XX (2013). We measured a significant intrinsic scatter of $\sigma_{\rm int} = 0.07 \pm 0.01$ dex. There is one spectacular outlier with an Y_{500}/Y_X ratio nearly twice as big as the mean. This is the *Planck* ESZ cluster identified with A2813 or RXC J0043.4-2037 in the REFLEX catalogue, located at $z = 0.29$. Its high ratio is very puzzling. It cannot result from an inaccurate redshift measurement, as this is based on spectroscopic data for several cluster galaxies (Böhringer et al. 2004). There is no evidence of a peculiar dynamical state from the X-ray morphology, and there is no evidence of contamination in the SZ data.

Part of the dispersion could be due to the use of an inappropriate fixed pressure profile in the Y_{500} extraction. When including possible errors on Y_{500} due to dispersion around the mean Arnaud et al. (2010) profile, the scatter is decreased to $\sigma_{\rm int} = 0.06$, a decrease at the $1\,\sigma$ level. To further assess the effect of the choice of the pressure profile, we re-extracted the SZ signal using the *Planck+XMM-Newton* profile shape measured for ESZ clusters by Planck Collaboration Int. V (2013, hereafter PIPV). Individual profiles are used for *Planck* ESZ clusters, and the mean profile is used for the other clusters. This should give the most reliable estimate of the Y_{500}–Y_X relation, since it is based directly on measured profile shapes. In this case, the slope and scatter remain unchanged but the normalization is slightly decreased (at the $0.5\,\sigma$ level). This is a result of the more inflated nature of the PIPV profile as compared to the Arnaud et al. (2010) REXCESS profile. The relation derived using PIPV pressure profiles is plotted in the right-hand panel of Fig. 32 together with the corresponding data points.

The relation does not exhibit significant evidence of variance of the Y_{500}/Y_X ratio with mass, the slope is consistent with unity, as expected from strong self-similarity of pressure profile shape. However, we found an intrinsic scatter about three times larger than the results of Kay et al. (2012). Partly this is due to the presence of outliers in our data set (as discussed above), or it may be due to projection effects in observed data sets (Kay et al. 2012). The mean ratio is very well constrained with a precision of 2.5%, $\log(Y_{500}/Y_X) = -0.027 \pm 0.010$. This confirms at higher precision the strong agreement between the SZ and X-ray measurements (within R_{500}) of the intra-cluster gas properties found by PEP XI and other studies (Andersson et al. 2011; Sifon et al. 2012; Marrone et al. 2012; Rozo et al. 2012). The ratio is perfectly consistent with the X-ray prediction, suggesting that the clumpiness must be low. However there are still large systematics that are discussed in Appendix D.

8. Summary

Planck's all-sky coverage and broad frequency range are designed to detect the SZ signal of galaxy clusters across the sky. We provide, from the first 15.5 months of observations, the largest ensemble of SZ-selected sources detected from an all-sky survey. The *Planck* catalogue of SZ sources contains 1227 detections. This catalogue, statistically characterized in terms of completeness and statistical reliability, was validated using external X-ray and optical/NIR data, alongside a multi-frequency follow-up programme for confirmation. A total of 861 SZ detections are confirmed associations with bona fide clusters, of which 178 are brand-new clusters. The remaining cluster candidates are divided into three classes according to their reliability, i.e., the quality of evidence that they are likely to be bona fide clusters.

A total of 813 *Planck* clusters have measured redshifts ranging from $z = 0.01$ to order one, with one-third of the clusters lying above $z = 0.3$. The brand-new *Planck* clusters extend the redshift range above $z = 0.3$. For all the *Planck* clusters with measured redshift, a mass can be estimated from the Compton Y measure. We provide a homogeneous mass estimate ranging from $(0.1 \text{ to } 1.6) \times 10^{15}\,\mathrm{M}_\odot$. Except at low redshifts, the *Planck* cluster distribution exhibits a nearly redshift-independent mass limit and occupies a unique region in the M–z space of massive, $M \geq 5 \times 10^{14}\,\mathrm{M}_\odot$, and high-redshift ($z \geq 0.5$) clusters. Owing to its all-sky nature, *Planck* detects new clusters in a region of the mass–redshift plane that is sparsely populated by the RASS catalogues. It detects the rarest clusters, i.e., the most massive clusters at high redshift in the exponential tail of the cluster mass function that are the most useful clusters for cosmological studies. With the presently confirmed *Planck* SZ detections, *Planck* doubles the number of massive clusters above redshift 0.5, as compared to other surveys. The *Planck* SZ catalogue is, and will be for years to come, the deepest all-sky SZ catalogue spanning the broadest cluster mass range.

The *Planck* SZ catalogue should motivate multi-wavelength follow-up efforts. The confirmation of the cluster candidates will reveal clusters at higher redshifts than the present distribution. Such follow-up efforts will further enhance the value of the *Planck* SZ catalogue as the first all-sky SZ selected catalogue. It will serve as a reference for studies of cluster physics (e.g., galaxy properties versus intra-cluster gas physics, dynamical state, evolution, etc.). Using an extended sub-sample of the *Planck* SZ clusters with high-quality *XMM-Newton* data, the scaling relations between SZ and X-ray properties were reassessed and updated. With better-quality data and thus higher precision, we show excellent agreement between SZ and X-ray measurements of the intra-cluster gas properties. We have thus derived a new up-to-date reference calibrated local relation between Y and Y_X.

The *Planck* SZ catalogue will also serve to define samples for cosmological studies. A first step in this direction is already taken in Planck Collaboration XX (2013), where an analysis of

the SZ cluster abundance to constrain the cosmological parameters is performed using a sub-sample selected from the PSZ catalogue consisting of 189 clusters detected above a signal-to-noise ratio of 7 with measured redshifts. The value-added information derived from the validation of the *Planck* SZ detections, in particular the SZ-based mass estimate, increases even further the value of the *Planck* SZ catalogue.

The combination of the *Planck* all-sky SZ data with near future and planned observations of the large-scale structure by surveys such as PAN-STARRS, LOFAR, Euclid, LSST, and RSG/e-ROSITA will revolutionize our understanding of large-scale structure formation and evolution.

Acknowledgements. The development of *Planck* has been supported by: ESA; CNES and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MICINN, JA and RES (Spain); Tekes, AoF and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); and PRACE (EU). A description of the *Planck* Collaboration and a list of its members, including the technical or scientific activities in which they have been involved, can be found at http://www.sciops.esa.int/index.php?project=planck&page=Planck_Collaboration. The authors thank N. Schartel, ESA *XMM-Newton* project scientist, for granting the Director Discretionary Time used for confirmation of SZ *Planck* candidates. The authors thank TUBITAK, IKI, KFU and AST for support in using RTT150 (Russian-Turkish 1.5-m telescope, Bakyrlytepe, Turkey); in particular we thank KFU and IKI for providing significant amount of their observing time at RTT150. We also thank BTA 6-m telescope Time Allocation Committee (TAC) for support of optical follow-up project. The authors acknowledge the use of the INT and WHT telescopes operated on the island of La Palma by the Isaac Newton Group of Telescopes at the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofisica de Canarias (IAC); the Nordic Optical Telescope, operated on La Palma jointly by Denmark, Finland, Iceland, Norway, and Sweden, at the Spanish Observatorio del Roque de los Muchachos of the IAC; the TNG telescope, operated on La Palma by the Fundacion Galileo Galilei of the INAF at the Spanish Observatorio del Roque de los Muchachos of the IAC; the GTC telescope, operated on La Palma by the IAC at the Spanish Observatorio del Roque de los Muchachos of the IAC; and the IAC80 telescope operated on the island of Tenerife by the IAC at the Spanish Observatorio del Teide of the IAC. Part of this research has been carried out with telescope time awarded by the CCI International Time Programme. The authors thank the TAC of the MPG/ESO-2.2m telescope for support of optical follow-up with WFI under *Max Planck* time. Observations were also conducted with ESO NTT at the La Silla Paranal Observatory. This research has made use of SDSS-III data. Funding for SDSS-III http://www.sdss3.org/ has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and DoE. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration.
This research has made use of the following databases: the NED and IRSA databases, operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the NASA; SIMBAD, operated at CDS, Strasbourg, France; SZ repository operated by IAS Data and Operation Center (IDOC) under contract with CNES and CNRS. The authors acknowledge the use of software provided by the US National Virtual Observatory.

References

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 715, 429
Abell, G. O. 1958, ApJS, 3, 211
Abell, G. O., Corwin, Jr., H. G., & Olowin, R. P. 1989, ApJS, 70, 1
Afanasiev, V. L. & Moiseev, A. V. 2005, Astronomy Letters, 31, 194
Aghanim, N., da Silva, A. C., & Nunes, N. J. 2009, A&A, 496, 637
Aghanim, N., de Luca, A., Bouchet, F. R., Gispert, R., & Puget, J. L. 1997, A&A, 325, 9
Aghanim, N. & Fromenteau, S. 2013, A&A
Aghanim, N., Hansen, S. H., & Lagache, G. 2005, A&A, 439, 901
Akritas, M. G. & Bershady, M. A. 1996, ApJ, 470, 706
Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, ARA&A, 49, 409
AMI Consortium, Hurley-Walker, N., Brown, M. L., et al. 2011, MNRAS, 414, L75
AMI Consortium, Zwart, J. T. L., Barker, R. W., et al. 2008, MNRAS, 391, 1545
Andersson, K., Benson, B. A., Ade, P. A. R., et al. 2011, ApJ, 738, 48
Angulo, R. E., Springel, V., White, S. D. M., et al. 2012, MNRAS, 426, 2046
Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2005, A&A, 441, 893
Arnaud, M., Pratt, G. W., Piffaretti, R., et al. 2010, A&A, 517, A92
Arnaud, M., XX, x., YY, y., & ZZ, z. 2013, A&A
Aussel, H., Pratt, G., & Arnaud, M. 2013, A&A
Bajkova, A. T. 2005, Astronomy Reports, 49, 947
Barbosa, D., Bartlett, J. G., Blanchard, A., & Oukbir, J. 1996, A&A, 314, 13
Belsole, E., Birkinshaw, M., & Worrall, D. M. 2005, MNRAS, 358, 120
Benítez, N. 2000, ApJ, 536, 571
Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393
Bock, D., Large, M. I., & Sadler, E. M. 1999, AJ, 117, 1578
Böhringer, H., Schuecker, P., Guzzo, L., et al. 2004, A&A, 425, 367
Böhringer, H., Schuecker, P., Pratt, G. W., et al. 2007, A&A, 469, 363
Böhringer, H., Voges, W., Huchra, J. P., et al. 2000, ApJS, 129, 435
Borgani, S. & Kravtsov, A. 2009, ArXiv e-prints
Borgani, S., Murante, G., Springel, V., et al. 2004, MNRAS, 348, 1078
Burenin, R. A., Vikhlinin, A., Hornstrup, A., et al. 2007, ApJS, 172, 561
Carvalho, P., Rocha, G., & Hobson, M. P. 2009, MNRAS, 393, 681
Carvalho, P., Rocha, G., Hobson, M. P., & Lasenby, A. 2011, ArXiv e-prints
Chon, G. & Böhringer, H. 2012, A&A, 538, A35
Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693
Crawford, C. S., Allen, S. W., Ebeling, H., Edge, A. C., & Fabian, A. C. 1999, MNRAS, 306, 857
Cunha, C. E., Lima, M., Oyaizu, H., Frieman, J., & Lin, H. 2009, MNRAS, 396, 2379
da Silva, A. C., Kay, S. T., Liddle, A. R., & Thomas, P. A. 2004, MNRAS, 348, 1401
Ebeling, H., Barrett, E., Donovan, D., et al. 2007, ApJ, 661, L33
Eckert, D., Molendi, S., & Paltani, S. 2011, A&A, 526, A79
Fabian, A. C. 2012, ARA&A, 50, 455
Fassbender, R., Böhringer, H., Nastasi, A., et al. 2011, New Journal of Physics, 13, 125014
Fromenteau, S., Bartlett, J., & Mei, S. 2013, in prep.
Giles, P. A., Maughan, B. J., Birkinshaw, M., Worrall, D. M., & Lancaster, K. 2012, MNRAS, 419, 503
Gioia, I. M., Henry, J. P., Mullis, C. R., et al. 2003, ApJS, 149, 29
Hao, J., McKay, T. A., Koester, B. P., et al. 2010, The Astrophysical Journal Supplement Series, 191, 254
Hasselfield, M., Hilton, M., Marriage, T. A., et al. 2013, ArXiv e-prints
Herranz, D., Sanz, J. L., Hobson, M. P., et al. 2002, MNRAS, 336, 1057
Hlavacek-Larrondo, J., Fabian, A. C., Edge, A. C., et al. 2012, MNRAS, 421, 1360
Hudson, D. S., Mittal, R., Reiprich, T. H., et al. 2010, A&A, 513, A37
Hurier, G., Hildebrandt, S. R., & Macias-Perez, J. F. 2010, ArXiv e-prints
Johnston, D. E., Sheldon, E. S., Wechsler, R. H., et al. 2007, ArXiv e-prints
Kay, S. T., Peel, M. W., Short, C. J., et al. 2012, MNRAS, 422, 1999
Kepner, J., Fan, X., Bahcall, N., et al. 1999, ApJ, 517, 78
Koester, B. P., McKay, T. A., Annis, J., et al. 2007, ApJ, 660, 239
Kravtsov, A. V., Vikhlinin, A., & Nagai, D. 2006, ApJ, 650, 128
Li, C. & White, S. 2013, A&A
Lin, Y., Partridge, B., Pober, J. C., et al. 2009, ApJ, 694, 992
Lin, Y.-T., Stanford, S. A., Eisenhardt, P. R. M., et al. 2012, ApJ, 745, L3
Magliocchetti, M. & Brüggen, M. 2007, MNRAS, 379, 260
Mann, A. W. & Ebeling, H. 2012, MNRAS, 420, 2120
Mantz, A., Allen, S. W., Ebeling, H., Rapetti, D., & Drlica-Wagner, A. 2010, MNRAS, 406, 1773
Marrone, D. P., Smith, G. P., Okabe, N., et al. 2012, ApJ, 754, 119
Maughan, B. J., Giles, P. A., Randall, S. W., Jones, C., & Forman, W. R. 2012, MNRAS, 421, 1583
McNamara, B. R. & Nulsen, P. E. J. 2012, New Journal of Physics, 14, 055023
Mehrtens, N., Romer, A. K., Hilton, M., et al. 2012, MNRAS, 423, 1024
Melin, J., Bartlett, J. G., & Delabrouille, J. 2006, A&A, 459, 341
Melin, J.-B., Aghanim, N., Bartelmann, M., et al. 2012, ArXiv e-prints
Menanteau, F., González, J., Juin, J., et al. 2010, ApJ, 723, 1523
Mitra, S., Rocha, G., Górski, K. M., et al. 2011, ApJS, 193, 5
Montier, L. A., Pelkonen, V., Juvela, M., Ristorcelli, I., & Marshall, D. J. 2010, A&A, 522, A83+
Motl, P. M., Hallman, E. J., Burns, J. O., & Norman, M. L. 2005, ApJ, 623, L63
Muchovej, S., Leitch, E., Culverhouse, T., Carpenter, J., & Sievers, J. 2012, ApJ, 749, 46
Nagai, D. 2006, ApJ, 650, 538
Nagai, D., Kravtsov, A. V., & Vikhlinin, A. 2007, ApJ, 668, 1
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Pacaud, F., Pierre, M., Adami, C., et al. 2007, MNRAS, 382, 1289
Peebles, P. J. E. 1980, The large-scale structure of the universe, ed. Peebles, P. J. E. (Princeton University Press)
Piffaretti, R., Arnaud, M., Pratt, G. W., Pointecouteau, E., & Melin, J.-B. 2011, A&A, 534, A109+
Piffaretti, R. & Valdarnini, R. 2008, A&A, 491, 71
Plagge, T., Benson, B. A., Ade, P. A. R., et al. 2010, ApJ, 716, 1118

Planck and AMI Collaborations. 2013, A&A, 550, A128
Planck Collaboration ES. 2013, The Explanatory Supplement to the Planck 2013 results (ESA)
Planck Collaboration I. 2013, In preparation
Planck Collaboration II. 2013, In preparation
Planck Collaboration Int. I. 2012, A&A, 543, A102
Planck Collaboration Int. III. 2013, A&A, 550, A129
Planck Collaboration Int. IV. 2013, A&A, 550, A130
Planck Collaboration Int. V. 2013, A&A, 550, A131
Planck Collaboration Int. VII. 2013, A&A, 550, A133
Planck Collaboration Int. X. 2013, Submitted to A&A
Planck Collaboration IX. 2011, A&A, 536, A9
Planck Collaboration VI. 2013, In preparation
Planck Collaboration VIII. 2011, A&A, 536, A8
Planck Collaboration X. 2011, A&A, 536, A10
Planck Collaboration XI. 2011, A&A, 536, A11
Planck Collaboration XII. 2013, In preparation
Planck Collaboration XIII. 2013, In preparation
Planck Collaboration XX. 2013, In preparation
Planck Collaboration XXI. 2013, In preparation
Planck Collaboration XXII. 2011, A&A, 536, A22
Planck Collaboration XXIII. 2011, A&A, 536, A23
Planck Collaboration XXVI. 2011, A&A, 536, A26
Planck Collaboration XXVIII. 2013, In preparation
Pratt, G. W., Croston, J. H., Arnaud, M., & Böhringer, H. 2009, A&A, 498, 361
Rafferty, D. A., McNamara, B. R., & Nulsen, P. E. J. 2008, ApJ, 687, 899
Rawle, T. D., Edge, A. C., Egami, E., et al. 2012, ApJ, 747, 29
Reichardt, C. L., Stalder, B., Bleem, L. E., et al. 2013, ApJ, 763, 127
Reichert, A., Böhringer, H., Fassbender, R., & Mühlegger, M. 2011, A&A, 535, A4
Reiprich, T. H. & Böhringer, H. 2002, ApJ, 567, 716
Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011, MNRAS, 410, 2481
Rozo, E., Rykoff, E. S., Evrard, A., et al. 2009, ApJ, 699, 768
Rozo, E., Vikhlinin, A., & More, S. 2012, ApJ, 760, 67
Rubiño-Martín, J. A. & Sunyaev, R. A. 2003, MNRAS, 344, 1155
Sayers, J., Czakon, N. G., Bridge, C., et al. 2012a, ApJ, 749, L15
Sayers, J., Czakon, N. G., Mantz, A., et al. 2012b, ArXiv e-prints
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schuecker, P., Böhringer, H., Collins, C. A., & Guzzo, L. 2003, A&A, 398, 867
SDSS-III Collaboration, :, Ahn, C. P., et al. 2012, ArXiv e-prints
Sifon, C., Menanteau, F., Hasselfield, M., et al. 2012, ArXiv e-prints
Story, K., Aird, K. A., Andersson, K., et al. 2011, ArXiv e-prints, arXiv:1102.2189
Sunyaev, R. A. & Zeldovich, I. B. 1980, ARA&A, 18, 537
Sunyaev, R. A. & Zeldovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173
Szabo, T., Pierpaoli, E., Dong, F., Pipino, A., & Gunn, J. 2011, ApJ, 736, 21
Takey, A., Schwope, A., & Lamer, G. 2011, A&A, 534, A120
Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709
Vanderlinde, K., Crawford, T. M., de Haan, T., et al. 2010, ApJ, 722, 1180
Vikhlinin, A., Allen, S. W., Arnaud, M., et al. 2009, in ArXiv Astrophysics e-prints, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey, 304–+
Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, ApJ, 640, 691
Vikhlinin, A., McNamara, B. R., Forman, W., et al. 1998, ApJ, 502, 558
Voges, W., Aschenbach, B., Boller, T., et al. 1999, A&A, 349, 389
Voges, W., Aschenbach, B., Boller, T., et al. 2000, VizieR Online Data Catalog, 9029, 0
Voit, G. M. 2005, Reviews of Modern Physics, 77, 207
Wen, Z. L., Han, J. L., & Liu, F. S. 2010, Monthly Notices of the Royal Astronomical Society, 407, 533
Wen, Z. L., Han, J. L., & Liu, F. S. 2012, ApJS, 199, 34
Wik, D. R., Sarazin, C. L., Ricker, P. M., & Randall, S. W. 2008, ApJ, 680, 17
Williamson, R., Benson, B. A., High, F. W., et al. 2011, ArXiv e-prints, arXiv:1101.1290
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579
Zwicky, F., Herzog, E., & Wild, P. 1961, Catalogue of galaxies and of clusters of galaxies, Vol. I

Appendix A: Selection of Frequency Channel Maps

An assessment of which combination of *Planck* frequency channels to use was performed using the MMF1 implementation of the matched multi-filter described in Sect. 2.2.2. The HFI and LFI channel maps were preprocessed as described in Sect. 2.1, with the only difference being that the point-source mask contained, in addition, detections from the LFI channel maps with S/N $\geq$ 10. Five different combinations of frequency channels were investigated, all *Planck* channels (30-857 GHz), all HFI channels plus the 70 GHz channel map from LFI (70-857 GHz), all HFI channels (100-857 GHz), the five lowest frequency HFI channels (100-545 GHz) and the four lowest frequency HFI channels (100-353 GHz). For each combination of frequency channels a catalogue of SZ sources was extracted, resulting in five different catalogues; the only differences between them must be entirely due to the choice of channels in the combination.

The first four of these catalogues are in good agreement in terms of the clusters detected, with all the differences amongst them being due to detections with S/N < 5. The (100-353 GHz) catalogue, however, contains significantly more detections, resulting in a poor agreement between it and the other catalogues that is not limited to low S/N detections. This is interpreted as being due to the lack of a dust-dominated channel in this combination, without which it is more difficult to constrain contamination due to dust emission.

In order to assess any improvement in the S/N ratios of detected clusters with the inclusion of extra data, a robust sample of reliable sources is required. To produce this, only clusters outside the 65% dust mask and with S/N $\geq$ 8 were kept from each combination. The differences in the S/N of the same sources detected using different frequency channel combinations can then be examined. The ratio between the S/N values of the common detections in each combination to those of the (100-857) combination was then found; the mean of this ratio is shown in Table A.1. This approach clearly shows the (100-353) combination to be considerably noisier than the other combinations, which is consistent with the observations reported above. Neither the inclusion of the LFI frequency channels or just the 70 GHz channel brings any significant improvement in the S/N of the clusters. Using the six HFI channel combination results in marginally better S/N than the (100-545 GHz) combination. The frequency channel combination chosen therefore is (100-857 GHz) since this gives the highest S/N with the smallest dataset. Reducing the S/N threshold from 8 to 6 and hence doubling the number of SZ sources used to evaluate the mean ratio does not change the conclusions of this analysis.

Appendix B: Extract from the *Planck* catalogue of SZ sources

We describe here the *Planck* catalogue of SZ sources delivered by the collaboration and available together with the individual lists from all three detections methods, the union mask used by these methods and the ensemble of notes on individual clusters at http://www.sciops.esa.int/index.php?page=Planck_Legacy_Archive&project=planck.

The union *Planck* SZ catalogue contains the coordinates and the signal-to-noise ratio of the SZ detections and a summary of the validation information, including external identification of the cluster and redshifts if they are available. The external identification quoted in the delivered product corresponds to the first identifier as defined in the external validation hierarchy, namely identification with MCXC clusters followed by Abell and Zwicky, followed by SDSS-based catalogues, followed by SZ catalogues, followed finally by searches in NED and SIMBAD. Due to the size–flux degeneracy discussed in

Table A.1: Mean of the ratio of the S/N of the common detections between the catalogues produced using different frequency channel combinations, excluding all detections inside the 35% dust mask. The improvement in the S/N of the detected clusters between the (100-353) and (100-857) combinations is clearly demonstrated, as is the lack of significant improvement in S/N when LFI data is included. The improvement between the (100-545) and (100-857) combinations is smaller and in the region of 1 to 2%.

Combination	(100-353)/(100-857)	(100-545)/(100-857)	(70-857)/(100-857)	(30-857)/(100-857)
S/N ≥ 6	0.86	0.99	1.00	1.00
S/N ≥ 8	0.84	0.98	1.00	1.00

Sect. 2.3, no reference flux quantity is outputted for the union catalogue.

The individual catalogues from the three detection methods, MMF1, MMF3, and PwS, contain the coordinates and the S/N ratio of the detections, and information on the size and flux of the clusters. The size is given in terms of θ_s and the flux is given in terms of the total integrated Comptonization parameter, $Y = Y_{5R_{500}}$. The full information on the degeneracy between and Y is provided in the form of the two-dimensional marginal probability distribution for each cluster.
The degeneracy information is provided in this form so that it can be combined with a model or external data to produce tighter constraints on the parameters. For example, combining it with an X-ray determination of the size can be done by taking a slice through the distribution at the appropriate θ_s. This is what is done in Sect. 7.2.1 and the refined measurement using X-ray information can be found in Table C.1.

Table B.1 presents an extract of the PSZ catalogue, in terms of the first rows of the online table and the following selected columns:
NAME: name of cluster.
R.A., DEC: right ascension (J2000) and declination (J2000).
S/N: signal-to-noise ratio of the SZ detection.
VALIDATION: status of the SZ detection from external validation: 20 = previously-known cluster; 10 = new confirmed *Planck* cluster; 1 = CLASS1 candidate ; 2 = CLASS2 candidate; 3 = CLASS3 candidate.
ID$_{\rm EXT}$: first external identifier of the known clusters.
z: redshift of the cluster as reported from the external validation.
COMMENTS: F = no comment; T = comment. Comments are readable in an external file.
The complete version of the PSZ catalogue also contains the additional columns:
INDEX: index of the detection, determined by the order of the clusters in the union catalogue and sorted into order of ascending Galactic longitude.
GLON, GLAT: Galactic coordinates.
POS_ERR: errors on the position.
PIPELINE: pipeline from which information is taken; namely 1 = MMF1; 2 = MMF3; 3 = PwS.
PIPE_DET: pipeline making the detection, with the following order in bits: 1st = MMF1; 2nd = MMF3; 3rd = PwS.
PCCS: flag for a match with sources from the PCCS catalogue.
COSMO: flag for those clusters that are included in the sample used for the cosmological analysis of Planck Collaboration XX (2013).

Appendix C: Outstanding outputs from the external validation

Based on the ancillary data used for the validation of the *Planck* SZ catalogue, we provide value-added information to the *Planck* SZ detections.
Namely, we provide, in addition to the first external identifier, possible other common identifiers, IDs.
We report the redshift information associated with the *Planck* clusters (z) and specify its source, (scr).
For clusters with measured redshifts, we compute the SZ-proxy Y_z and the mass estimate ($M_{500}^{Y_z}$) and associated errors. For the clusters identified with MCXC clusters we provide the SZ signal, $Y_{500,\rm PSX}$, re-extracted fixing the size to the X-ray size provided in the MCXC catalogue at the X-ray position. Note that the X-ray positions used in the present study are those quoted in the MCXC meta-catalogue. The positions reported in the ESZ sample were taken from a sampled grid of coordinates with a pixel size of 1.71 arcmin. Due to this sampling, the reported MCXC positions in the ESZ sample exhibit an average offset of 70 arcsec (less than a pixel, which varies depending on the position of the object on the sphere).
For *Planck* SZ detections not associated with a previously-known X-ray cluster and with a signal-to-noise, $(\rm S/N)_{RASS} \geq 1\,\sigma$, we provide the unabsorbed X-ray flux, $S_{\rm X}$ (and error), measured in an aperture of 5 arcmin in the band [0.1-2.4] keV. We only provide an upper limit in the case of $(\rm S/N)_{RASS} < 1\,\sigma$, except for three SZ detections for which RASS exposure is very low and $(\rm S/N)_{RASS} < -5\,\sigma$. The aperture is centred on the *Planck* position, except for candidates associated with a BSC source for which we adopt the X-ray position. These clusters are flagged.

Appendix D: Systematic effects on the X-ray versus SZ scaling relation

Both X-ray and SZ measurements are likely affected by systematic effects linked to e.g., background estimation and subtraction methods, calibration issues, etc. One sign of the impact of these effects is the fact that the slope of the relation between Y_{500} flux and $Y_{\rm X}/D_{\rm A}^2$ in units of arcmin^2 is $\alpha = 0.91 \pm 0.02$, which is significantly smaller than unity even after Malmquist bias correction. As this is not the case for the relation in physical units (Mpc^2), the observed slope cannot be due to a true physical variation in the ratio (e.g., with mass).

SZ fluxes are subject to uncertainties due to systematic differences between measurement methods. From the comparison between PwS and MMF photometry (Sect. 2.3), we estimate that the net effect is typically 0.03 dex. The effect is independent of SZ flux, thus cannot explain the shallower than expected slope.

Uncertainties in the X-ray measurements are dominated by temperature uncertainties due to calibration systematics. We can investigate the magnitude of these effects by examining the re-

Table B.1: Extract from the *Planck* catalogue of SZ sources. First rows of the online table are shown. The online table contains additional columns as documented in the Explanatory Supplement, and in the text.

NAME	R.A. degrees	Dec. degrees	S/N	Validation	ID_{EXT}	z	Comments
PSZ1 G000.08+45.15	229.19790	-0.9792795	4.60	20	RXC J1516.5-0056	0.1198	F
PSZ1 G000.42-41.84	316.06990	-41.339730	5.99	20	RXC J2104.3-4120	0.1651	F
PSZ1 G000.42-41.84	307.93571	-40.595198	5.30	20	RXC J2031.8-4037	0.3416	F
PSZ1 G000.77-35.67	244.58411	-13.070074	6.04	3	...	...	F
PSZ1 G001.00+25.71	349.60728	-36.278003	4.50	20	ACO S 1109	0.1400	F
PSZ1 G002.24-68.27	334.65975	-38.880540	7.84	20	RXC J2218.6-3853	0.1411	F
PSZ1 G002.77-56.16	234.99997	-3.2929395	7.03	20	RXC J1540.1-0318	0.1533	F
PSZ1 G002.80+39.24	292.16440	-35.711064	4.92	3	...	...	F

Table C.1: Additional information from the external validation process. `IDs` represents the first external identifier and possible other common identifiers. z is the cluster redshift. `scr` is the source from which redshift was taken. The `scr` takes defined values, e.g., 11 for redshifts from the MCXC meta-catalogue Piffaretti et al. (2011), 17 for redshifts taken from NED or SIMBAD databases, etc. Y_z is the SZ-proxy with asymmetric errors. $M_{500}^{Y_z}$ is the derived mass estimate with asymmetric errors. $Y_{500,PSX}$ is the SZ signal, re-extracted fixing the size to the X-ray size provided in the MCXC catalogue at the X-ray position, for PSZ clusters identified with MCXC clusters. S_X is the unabsorbed X-ray flux measured in an aperture of 5 arcmin in the band [0.1-2.4] keV. The aperture is centred on the *Planck* position, except for candidates associated with a BSC source for which we adopt the X-ray position. For sources with $(S/N)_{RASS} < 1\,\sigma$, we only quote an upper limit. The full table will be available at `http://www.sciops.esa.int/index.php?page=Planck_Legacy_Archive&project=planck.` upon final acceptance of the article. Until then, it can be provided upon request.

NAME	z (scr)	Y_z	$M_{500}^{Y_z}$	S_X	IDs	$Y_{500,PSX}$
PSZ1 G000.08+45.15	0.1198 (11)	$12.35^{3.43}_{3.33}$	$3.10^{0.45}_{0.50}$	...	RXC J1516.5-0056, A2051	2.73
PSZ1 G000.42-41.84	0.1651 (11)	$14.05^{2.78}_{2.70}$	$4.46^{0.47}_{0.50}$	...	RXC J2104.3-4120, A3739	1.28
PSZ1 G000.42-41.84	0.3416 (11)	$9.14^{1.98}_{1.93}$	$6.20^{0.72}_{0.77}$	...	RXC J2031.8-4037	11.80
PSZ1 G000.77-35.67	... (-1)	...	...	≤ 1.35	...	...
PSZ1 G001.00+25.71	0.1400 (17)	$7.43^{2.71}_{2.61}$	$2.69^{0.51}_{0.58}$	1.74 ± 0.65	ACO S 1109	...
PSZ1 G002.24-68.27	0.1411 (11)	$18.29^{2.92}_{2.85}$	$4.49^{0.39}_{0.41}$	...	RXC J2218.6-3853, A3856	7.94
PSZ1 G002.77-56.16	0.1533 (11)	$26.14^{4.68}_{4.53}$	$5.91^{0.57}_{0.60}$	...	RXC J1540.1-0318, A2104	0.20
PSZ1 G002.80+39.24	... (-1)	...	...	≤ -0.07	...	...

lation between the Y_X values obtained with *XMM-Newton* by Planck Collaboration XI (2011, hereafter the PEP XI ESZ-XMM sample) to those obtained with *Chandra* in a study of 28 clusters from the same sample by Rozo et al. (2012) (hereafter the ESZ–*Chandra* sample). The *Chandra* values are larger, with a mean offset of 0.02 dex. However, there is no significant evidence of variation with Y_X, thus X-ray calibration issues again cannot explain the observed slope.

A further source of uncertainty in X-ray measurements concerns the X-ray analysis method (e.g., due to background estimation and subtraction of point sources and substructure). Rozo et al. (2012) noted the difference between the ratio obtained with ESZ–*Chandra* and ESZ-XMM samples and suggested that it might be due to *XMM-Newton* data analysis issues. The PEP XI ESZ-XMM sample was analyzed by two independent methods depending on the cluster extension in the field-of-view. Sub-sample A consisted of 19 nearby clusters that extend beyond the *XMM-Newton* field–of–view, and for which direct background estimates are not possible, while the background for the remaining 43 objects was estimated using a region external to the cluster. The ESZ–*Chandra* sample studied by Rozo et al. (2012) consists mostly sub-sample A objects. While systematic effects due to background estimation are certainly more important for sub-sample A than for sub-sample B, these effects cannot fully explain the observed behaviour of the Y_{500}/Y_X ratio. Indeed, excluding sub-sample A clusters, the slope of the Y_{500}–Y_X/D_A^2 relation is $\alpha = 0.89 \pm 0.04$, still significantly smaller than unity. The origin of the systematic differences between sub-sample A and B objects is unclear.

The variation of the Y_{500}/Y_X ratio with flux remains largely unexplained. It may be due to residual Malmquist bias, in addition to a complex combination of systematic effects in SZ and X–ray measurements. For instance, we note that higher flux clusters correspond to nearby objects that have larger angular sizes. The background estimate in both X-ray and SZ signals is subject to larger uncertainty in this case.

The lack of a complete explanation for the observed slope of the Y_{500}–Y_X relation, and its ultimate correction, has several implications. Firstly, the shallower slope in units of arcmin^2 translates into an over-estimate of the dispersion about the relation when measured in Mpc^2. From the difference in intrinsic scatter about the relation in both physical and arcmin units, we estimate that this effect contributes at the level of about 0.01 dex to the scatter seen in the physical Y_{500}–Y_X relation.

Secondly, the Y_{500}/Y_X ratio will depend on the exact sample definition, via the range of fluxes probed. The observed slope of $\alpha = 0.91 \pm 0.02$ translates into a variation of about ± 0.06 dex of the Y_{500}/Y_X ratio over the range of SZ fluxes studied here. The

ESZ–*Chandra* objects studied by Rozo et al. (2012) lie preferentially at high fluxes, with a median flux two times higher than the PEP XI-XMM sample. For $\alpha = 0.91$, this will translate into a roughly 0.03 dex difference in the Y_{500}/Y_X ratio. The Y_{500}/Y_X ratio found by Rozo et al. (2012), $\log(Y_{500}/Y_X) = -0.088 \pm 0.012$, is significantly lower than our value of -0.027 ± 0.010. However, it can be explained by a combination of their sample definition, a neglect of Malmquist bias, and the aforementioned calibration issues between *XMM-Newton* and *Chandra*.

In summary, uncertainties on the Y_{500}/Y_X ratio are dominated by systematic effects in both X–ray and SZ measurements. This unfortunately precludes any definitive statement on the magnitude of the gas clumpiness within R_{500}. Follow-up of well-defined sub-samples (e.g., above a given S/N) should help to disentangle biases due to sample selection and measurement of the different quantities.

[1] APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/lrfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France
[2] Aalto University Metsähovi Radio Observatory, Metsähovintie 114, FIN-02540 Kylmälä, Finland
[3] Academy of Sciences of Tatarstan, Bauman Str., 20, Kazan, 420111, Republic of Tatarstan, Russia
[4] African Institute for Mathematical Sciences, 6-8 Melrose Road, Muizenberg, Cape Town, South Africa
[5] Agenzia Spaziale Italiana Science Data Center, c/o ESRIN, via Galileo Galilei, Frascati, Italy
[6] Agenzia Spaziale Italiana, Viale Liegi 26, Roma, Italy
[7] Astrophysics Group, Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, U.K.
[8] Astrophysics & Cosmology Research Unit, School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Private Bag X54001, Durban 4000, South Africa
[9] Atacama Large Millimeter/submillimeter Array, ALMA Santiago Central Offices, Alonso de Cordova 3107, Vitacura, Casilla 763 0355, Santiago, Chile
[10] CITA, University of Toronto, 60 St. George St., Toronto, ON M5S 3H8, Canada
[11] CNRS, IRAP, 9 Av. colonel Roche, BP 44346, F-31028 Toulouse cedex 4, France
[12] California Institute of Technology, Pasadena, California, U.S.A.
[13] Centre for Theoretical Cosmology, DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA U.K.
[14] Centro de Astrofísica, Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
[15] Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Plaza San Juan, 1, planta 2, E-44001, Teruel, Spain
[16] Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.
[17] Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain
[18] DSM/Irfu/SPP, CEA-Saclay, F-91191 Gif-sur-Yvette Cedex, France
[19] DTU Space, National Space Institute, Technical University of Denmark, Elektrovej 327, DK-2800 Kgs. Lyngby, Denmark
[20] Département de Physique Théorique, Université de Genève, 24, Quai E. Ansermet,1211 Genève 4, Switzerland
[21] Departamento de Física Fundamental, Facultad de Ciencias, Universidad de Salamanca, 37008 Salamanca, Spain
[22] Departamento de Física, Universidad de Oviedo, Avda. Calvo Sotelo s/n, Oviedo, Spain
[23] Department of Astronomy and Astrophysics, University of Toronto, 50 Saint George Street, Toronto, Ontario, Canada
[24] Department of Astronomy and Geodesy, Kazan Federal University, Kremlevskaya Str., 18, Kazan, 420008, Russia
[25] Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
[26] Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California, U.S.A.
[27] Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia, Canada
[28] Department of Physics and Astronomy, Dana and David Dornsife College of Letter, Arts and Sciences, University of Southern California, Los Angeles, CA 90089, U.S.A.
[29] Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.
[30] Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, U.K.
[31] Department of Physics, Gustaf Hällströmin katu 2a, University of Helsinki, Helsinki, Finland
[32] Department of Physics, Princeton University, Princeton, New Jersey, U.S.A.
[33] Department of Physics, University of California, Berkeley, California, U.S.A.
[34] Department of Physics, University of California, One Shields Avenue, Davis, California, U.S.A.
[35] Department of Physics, University of California, Santa Barbara, California, U.S.A.
[36] Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois, U.S.A.
[37] Dipartimento di Fisica e Astronomia G. Galilei, Università degli Studi di Padova, via Marzolo 8, 35131 Padova, Italy
[38] Dipartimento di Fisica e Scienze della Terra, Università di Ferrara, Via Saragat 1, 44122 Ferrara, Italy
[39] Dipartimento di Fisica, Università La Sapienza, P. le A. Moro 2, Roma, Italy
[40] Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria, 16, Milano, Italy
[41] Dipartimento di Fisica, Università degli Studi di Trieste, via A. Valerio 2, Trieste, Italy
[42] Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy
[43] Discovery Center, Niels Bohr Institute, Blegdamsvej 17, Copenhagen, Denmark
[44] Dpto. Astrofísica, Universidad de La Laguna (ULL), E-38206 La Laguna, Tenerife, Spain
[45] European Southern Observatory, ESO Vitacura, Alonso de Cordova 3107, Vitacura, Casilla 19001, Santiago, Chile
[46] European Space Agency, ESAC, Planck Science Office, Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, Madrid, Spain
[47] European Space Agency, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[48] Finnish Centre for Astronomy with ESO (FINCA), University of Turku, Väisäläntie 20, FIN-21500, Piikkiö, Finland
[49] GEPI, Observatoire de Paris, Section de Meudon, 5 Place J. Janssen, 92195 Meudon Cedex, France
[50] Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
[51] INAF - Osservatorio Astrofisico di Catania, Via S. Sofia 78, Catania, Italy
[52] INAF - Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, Padova, Italy
[53] INAF - Osservatorio Astronomico di Roma, via di Frascati 33, Monte Porzio Catone, Italy
[54] INAF - Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, Trieste, Italy
[55] INAF/IASF Bologna, Via Gobetti 101, Bologna, Italy
[56] INAF/IASF Milano, Via E. Bassini 15, Milano, Italy
[57] INFN, Sezione di Bologna, Via Irnerio 46, I-40126, Bologna, Italy
[58] INFN, Sezione di Roma 1, Università di Roma Sapienza, Piazzale Aldo Moro 2, 00185, Roma, Italy

[59] IPAG: Institut de Planétologie et d'Astrophysique de Grenoble, Université Joseph Fourier, Grenoble 1 / CNRS-INSU, UMR 5274, Grenoble, F-38041, France
[60] ISDC Data Centre for Astrophysics, University of Geneva, ch. d'Ecogia 16, Versoix, Switzerland
[61] IUCAA, Post Bag 4, Ganeshkhind, Pune University Campus, Pune 411 007, India
[62] Imperial College London, Astrophysics group, Blackett Laboratory, Prince Consort Road, London, SW7 2AZ, U.K.
[63] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, U.S.A.
[64] Institut Néel, CNRS, Université Joseph Fourier Grenoble I, 25 rue des Martyrs, Grenoble, France
[65] Institut Universitaire de France, 103, bd Saint-Michel, 75005, Paris, France
[66] Institut d'Astrophysique Spatiale, CNRS (UMR8617) Université Paris-Sud 11, Bâtiment 121, Orsay, France
[67] Institut d'Astrophysique de Paris, CNRS (UMR7095), 98 bis Boulevard Arago, F-75014, Paris, France
[68] Institute for Space Sciences, Bucharest-Magurale, Romania
[69] Institute of Astronomy and Astrophysics, Academia Sinica, Taipei, Taiwan
[70] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, U.K.
[71] Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway
[72] Instituto de Astrofísica de Canarias, C/Vía Láctea s/n, La Laguna, Tenerife, Spain
[73] Instituto de Física de Cantabria (CSIC-Universidad de Cantabria), Avda. de los Castros s/n, Santander, Spain
[74] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California, U.S.A.
[75] Jodrell Bank Centre for Astrophysics, Alan Turing Building, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester, M13 9PL, U.K.
[76] Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 0HA, U.K.
[77] LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France
[78] LERMA, CNRS, Observatoire de Paris, 61 Avenue de l'Observatoire, Paris, France
[79] Laboratoire AIM, IRFU/Service d'Astrophysique - CEA/DSM - CNRS - Université Paris Diderot, Bât. 709, CEA-Saclay, F-91191 Gif-sur-Yvette Cedex, France
[80] Laboratoire Traitement et Communication de l'Information, CNRS (UMR 5141) and Télécom ParisTech, 46 rue Barrault F-75634 Paris Cedex 13, France
[81] Laboratoire de Physique Subatomique et de Cosmologie, Université Joseph Fourier Grenoble I, CNRS/IN2P3, Institut National Polytechnique de Grenoble, 53 rue des Martyrs, 38026 Grenoble cedex, France
[82] Laboratoire de Physique Théorique, Université Paris-Sud 11 & CNRS, Bâtiment 210, 91405 Orsay, France
[83] Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.
[84] MPA Partner Group, Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Chinese Academy of Sciences, Nandan Road 80, Shanghai 200030, China
[85] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
[86] Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstraße, 85748 Garching, Germany
[87] McGill Physics, Ernest Rutherford Physics Building, McGill University, 3600 rue University, Montréal, QC, H3A 2T8, Canada
[88] MilliLab, VTT Technical Research Centre of Finland, Tietotie 3, Espoo, Finland
[89] Moscow Institute of Physics and Technology, Dolgoprudny, Institutsky per., 9, 141700, Russia
[90] Niels Bohr Institute, Blegdamsvej 17, Copenhagen, Denmark
[91] Observational Cosmology, Mail Stop 367-17, California Institute of Technology, Pasadena, CA, 91125, U.S.A.
[92] Optical Science Laboratory, University College London, Gower Street, London, U.K.
[93] SB-ITP-LPPC, EPFL, CH-1015, Lausanne, Switzerland
[94] SISSA, Astrophysics Sector, via Bonomea 265, 34136, Trieste, Italy
[95] SUPA, Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, U.K.
[96] School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff, CF24 3AA, U.K.
[97] Space Research Institute (IKI), Russian Academy of Sciences, Profsoyuznaya Str, 84/32, Moscow, 117997, Russia
[98] Space Sciences Laboratory, University of California, Berkeley, California, U.S.A.
[99] Special Astrophysical Observatory, Russian Academy of Sciences, Nizhnij Arkhyz, Zelenchukskiy region, Karachai-Cherkessian Republic, 369167, Russia
[100] Stanford University, Dept of Physics, Varian Physics Bldg, 382 Via Pueblo Mall, Stanford, California, U.S.A.
[101] Sub-Department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, U.K.
[102] TÜBİTAK National Observatory, Akdeniz University Campus, 07058, Antalya, Turkey
[103] Theory Division, PH-TH, CERN, CH-1211, Geneva 23, Switzerland
[104] UPMC Univ Paris 06, UMR7095, 98 bis Boulevard Arago, F-75014, Paris, France
[105] Universität Heidelberg, Institut für Theoretische Astrophysik, Philosophenweg 12, 69120 Heidelberg, Germany
[106] Université Denis Diderot (Paris 7), 75205 Paris Cedex 13, France
[107] Université de Toulouse, UPS-OMP, IRAP, F-31028 Toulouse cedex 4, France
[108] Universities Space Research Association, Stratospheric Observatory for Infrared Astronomy, MS 232-11, Moffett Field, CA 94035, U.S.A.
[109] University Observatory, Ludwig Maximilian University of Munich, Scheinerstrasse 1, 81679 Munich, Germany
[110] University of Granada, Departamento de Física Teórica y del Cosmos, Facultad de Ciencias, Granada, Spain
[111] University of Miami, Knight Physics Building, 1320 Campo Sano Dr., Coral Gables, Florida, U.S.A.
[112] Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland

Paper II

***Planck* 2013 results. XII. Component separation**

Planck Collaboration

Submitted to Astronomy & Astrophysics

Astronomy & Astrophysics manuscript no. component˙separation
March 22, 2013

Planck 2013 results. XII. Component separation

Planck Collaboration: P. A. R. Ade[90], N. Aghanim[63], C. Armitage-Caplan[95], M. Arnaud[77], M. Ashdown[74,6*], F. Atrio-Barandela[20], J. Aumont[63], C. Baccigalupi[89], A. J. Banday[98,11], R. B. Barreiro[70], J. G. Bartlett[1,72], E. Battaner[100], K. Benabed[64,97], A. Benoît[61], A. Benoit-Lévy[28,64,97], J.-P. Bernard[11], M. Bersanelli[37,52], P. Bielewicz[98,11,89], J. Bobin[77], J. J. Bock[72,12], A. Bonaldi[73], L. Bonavera[70], J. R. Bond[9], J. Borrill[15,93], F. R. Bouchet[64,97], F. Boulanger[63], M. Bridges[74,6,67], M. Bucher[1], C. Burigana[51,35], R. C. Butler[51], J.-F. Cardoso[78,1,64], A. Catalano[79,76], A. Challinor[67,74,13], A. Chamballu[77,17,63], R.-R. Chary[60], X. Chen[60], L.-Y Chiang[66], H. C. Chiang[30,7], P. R. Christensen[85,40], S. Church[94], D. L. Clements[59], S. Colombi[64,97], L. P. L. Colombo[27,72], F. Couchot[75], A. Coulais[76], B. P. Crill[72,86], M. Cruz[22], A. Curto[6,70], F. Cuttaia[51], L. Danese[89], R. D. Davies[73], R. J. Davis[73], P. de Bernardis[36], A. de Rosa[51], G. de Zotti[48,89], J. Delabrouille[1], J.-M. Delouis[64,97], F.-X. Désert[56], C. Dickinson[73], J. M. Diego[70], H. Dole[63,62], S. Donzelli[52], O. Doré[72,12], M. Douspis[63], J. Dunkley[95], X. Dupac[43], G. Efstathiou[67], T. A. Enßlin[82], H. K. Eriksen[68], E. Falgarone[76], F. Finelli[51,53], O. Forni[98,11], M. Frailis[50], A. A. Fraisse[30], E. Franceschi[51], S. Galeotta[50], K. Ganga[1], M. Giard[98,11], G. Giardino[44], Y. Giraud-Héraud[1], J. González-Nuevo[70,89], K. M. Górski[72,102], S. Gratton[74,67], A. Gregorio[38,50], A. Gruppuso[51], F. K. Hansen[68], D. Hanson[83,72,9], D. Harrison[67,74], G. Helou[12], S. Henrot-Versillé[75], C. Hernández-Monteagudo[14,82], D. Herranz[70], S. R. Hildebrandt[12], E. Hivon[64,97], M. Hobson[6], W. A. Holmes[72], A. Hornstrup[18], W. Hovest[82], G. Huey[33], K. M. Huffenberger[101], T. R. Jaffe[98,11], A. H. Jaffe[59], J. Jewell[72], W. C. Jones[30], M. Juvela[29], E. Keihänen[29], R. Keskitalo[25,15], T. S. Kisner[81], R. Kneissl[42,8], J. Knoche[82], L. Knox[31], M. Kunz[19,63,3], H. Kurki-Suonio[29,46], G. Lagache[63], A. Lähteenmäki[2,46], J.-M. Lamarre[76], A. Lasenby[6,74], R. J. Laureijs[44], C. R. Lawrence[72], M. Le Jeune[1], S. Leach[89], J. P. Leahy[73], R. Leonardi[43], J. Lesgourgues[96,88], M. Liguori[34], P. B. Lilje[68], M. Linden-Vørnle[18], M. López-Caniego[70], P. M. Lubin[32], J. F. Macías-Pérez[79], B. Maffei[73], D. Maino[37,52], N. Mandolesi[51,5,35], A. Marcos-Caballero[70], M. Maris[50], D. J. Marshall[77], P. G. Martin[9], E. Martínez-González[70], S. Masi[36], S. Matarrese[34], F. Matthai[82], P. Mazzotta[39], P. R. Meinhold[32], A. Melchiorri[36,54], L. Mendes[43], A. Mennella[37,52], M. Migliaccio[67,74], K. Mikkelsen[68], S. Mitra[58,72], M.-A. Miville-Deschênes[63,9], A. Moneti[64], L. Montier[98,11], G. Morgante[51], D. Mortlock[59], A. Moss[91], D. Munshi[90], P. Naselsky[85,40], F. Nati[36], P. Natoli[35,4,51], C. B. Netterfield[23], H. U. Nørgaard-Nielsen[18], F. Noviello[73], D. Novikov[59], I. Novikov[85], I. J. O'Dwyer[72], S. Osborne[94], C. A. Oxborrow[18], F. Paci[89], L. Pagano[36,54], F. Pajot[63], R. Paladini[60], D. Paoletti[51,53], B. Partridge[45], F. Pasian[50], G. Patanchon[1], T. J. Pearson[12,60], O. Perdereau[75], L. Perotto[79], F. Perrotta[89], V. Pettorino[19], F. Piacentini[36], M. Piat[1], E. Pierpaoli[27], D. Pietrobon[72], S. Plaszczynski[75], P. Platania[71], E. Pointecouteau[98,11], G. Polenta[4,49], N. Ponthieu[63,56], L. Popa[65], T. Poutanen[46,29,2], G. W. Pratt[77], G. Prézeau[12,72], S. Prunet[64,97], J.-L. Puget[63], J. P. Rachen[24,82], W. T. Reach[99], R. Rebolo[69,16,41], M. Reinecke[82], M. Remazeilles[63,1], C. Renault[79], A. Renzi[89], S. Ricciardi[51], T. Riller[82], I. Ristorcelli[98,11], G. Rocha[72,12], C. Rosset[1], G. Roudier[1,76,72], M. Rowan-Robinson[59], J. A. Rubiño-Martín[69,41], B. Rusholme[60], E. Salerno[10], M. Sandri[51], D. Santos[79], G. Savini[87], F. Schiavon[51], D. Scott[26], M. D. Seiffert[72,12], E. P. S. Shellard[13], L. D. Spencer[90], J.-L. Starck[77], R. Stompor[1], R. Sudiwala[90], R. Sunyaev[82,92], F. Sureau[77], D. Sutton[67,74], A.-S. Suur-Uski[29,46], J.-F. Sygnet[64], J. A. Tauber[44], D. Tavagnacco[50,38], L. Terenzi[51], L. Toffolatti[21,70], M. Tomasi[52], M. Tristram[75], M. Tucci[19,75], J. Tuovinen[84], M. Türler[57], G. Umana[47], L. Valenziano[51], J. Valiviita[46,29,68], B. Van Tent[80], J. Varis[84], M. Viel[50,55], P. Vielva[70], F. Villa[51], N. Vittorio[39], L. A. Wade[72], B. D. Wandelt[64,97,33], I. K. Wehus[72], A. Wilkinson[73], J.-Q. Xia[89], D. Yvon[17], A. Zacchei[50], and A. Zonca[32]

(Affiliations can be found after the references)

Preprint online version: March 22, 2013

ABSTRACT

Planck has produced detailed all-sky observations over nine frequency bands between 30 and 857 GHz. These observations allow robust reconstruction of the primordial cosmic microwave background (CMB) temperature fluctuations over nearly the full sky, as well as new constraints on Galactic foregrounds, including thermal dust and line emission from molecular carbon monoxide (CO). This paper describes the component separation framework adopted by *Planck* for many cosmological analyses, including CMB power spectrum determination and likelihood construction on large angular scales, studies of primordial non-Gaussianity and statistical isotropy, the integrated Sachs-Wolfe effect (ISW), and gravitational lensing, and searches for topological defects. We test four foreground-cleaned CMB maps derived using qualitatively different component separation algorithms. The quality of our reconstructions is evaluated through detailed simulations and internal comparisons, and shown through various tests to be internally consistent and robust for CMB power spectrum and cosmological parameter estimation up to $\ell = 2000$. The parameter constraints on ΛCDM cosmologies derived from these maps are consistent with those presented in the cross-spectrum based *Planck* likelihood analysis. We choose two of the CMB maps for specific scientific goals. We also present maps and frequency spectra of the Galactic low-frequency, CO, and thermal dust emission. The component maps are found to provide a faithful representation of the sky, as evaluated by simulations, with the largest bias seen in the CO component at 3 %. For the low-frequency component, the spectral index varies widely over the sky, ranging from about $\beta = -4$ to -2. Considering both morphology and prior knowledge of the low frequency components, the index map allows us to associate a steep spectral index ($\beta < -3.2$) with strong anomalous microwave emission, corresponding to a spinning dust spectrum peaking below 20 GHz, a flat index of $\beta > -2.3$ with strong free-free emission, and intermediate values with synchrotron emission.

Key words. cosmology: observations

* Corresponding author: M. Ashdown, maja1@mrao.cam.ac.uk.

1. Introduction

This paper, one of a set associated with the 2013 release of data from the *Planck*[1] mission (Planck Collaboration I 2013), describes the component separation techniques applied to the *Planck* data to produce maps of the cosmic microwave background (CMB) temperature anisotropies (see Fig. 1) and of diffuse foregrounds.

The sky at millimetre and sub-millimetre wavelengths contains a wealth of cosmological and astrophysical information. Accessing it is an inversion process, known as component separation, to extract the sources of emission contributing to a set of maps observed at different frequencies. *Planck* gives us a powerful data set to unlock new information in this manner by observing the entire sky from 30 to 857 GHz in nine frequency bands at higher angular resolution and sensitivity than its predecessors. Accurate and detailed component separation is a central objective of the mission.

We divide the foregrounds into two distinct categories: diffuse emission from the Galaxy and compact sources. The Galactic foregrounds are the principal source of contamination of the CMB on large angular scales, with fluctuation power decreasing roughly as a power law towards higher multipoles (Bennett et al. 2003). They are dominated by synchrotron, free-free and Anomalous Microwave Emission (AME, ascribed to spinning dust grains, at frequencies below 70 GHz) and by rotational line emission from carbon monoxide (CO) molecules and thermal dust emission at frequencies above 100 GHz. Extragalactic foregrounds, on the other hand, dominate the small-scale contamination of the CMB. They arise from discrete, individually detectable compact sources and the collective emission from unresolved radio and infrared (IR) sources, and also from the Sunyaev-Zeldovich (SZ) effect in galaxy clusters (Planck Collaboration XVIII 2011; Planck Collaboration XXVIII 2013; Planck Collaboration XXIX 2013).

In the *Planck* analyses, these foregrounds are dealt with in a variety of ways. At the power spectrum and likelihood level, the extragalactic foregrounds are modelled with parameterized power spectra, appropriate to their statistical isotropy, over regions restricted to low Galactic emission (Planck Collaboration XV 2013). Component separation as described in the present paper aims at removing Galactic emission to produce CMB maps covering the largest possible sky area for studies of the large-scale properties and higher-order statistics of the CMB. In addition, this component separation provides a reconstruction of the diffuse emission from our Galaxy. Detailed studies of specific extragalactic foregrounds, such as the cosmic infrared background (CIB) (Planck Collaboration XVIII 2013) and the diffuse Sunyaev-Zeldovich (SZ) signal (Planck Collaboration XXI 2013), employ methods tailored to their particular needs.

Building on previous work (Leach et al. 2008), we approach CMB extraction with a philosophy designed to ensure robustness by applying four distinct algorithms based on two different methodologies. The first avoids any assumptions concerning the foregrounds and relies solely on a minimum variance criterion for the data component possessing a blackbody spectrum (i.e., the CMB), while the second methodology relies on parametric modelling of the foregrounds in either real or harmonic space. We evaluate the performance of these component separation algorithms through detailed simulations, and we examine the robustness of the recovered CMB maps by comparing them, their power spectra, and their resulting cosmological constraints. As a diagnostic, we also briefly examine their higher-order statistics.

The CMB results presented in this work serve a number applications. We use the real-space modelling to produce a clean CMB map and power spectra on large angular scales, where diffuse Galactic emission is the main contaminant, to construct the likelihood function at low multipoles; this is then combined with the high multipole likelihood function that models extragalactic foregrounds with power spectra (Planck Collaboration XV 2013). The high resolution CMB maps are used as a check on primary cosmological constraints (see below), for lensing studies (Planck Collaboration XVII 2013), studies of the integrated Sachs-Wolfe effect (Planck Collaboration XIX 2013), of the isotropy of the CMB (Planck Collaboration XXIII 2013), of non-Gaussian statistics (Planck Collaboration XXIV 2013), in searches for topological defects (Planck Collaboration XXV 2013), and for examination of the geometry and topology of the Universe (Planck Collaboration XXVI 2013).

In addition, we present maps of diffuse Galactic emission divided into low- and high-frequency components, as well as a molecular CO component. We judge the adequacy of this reconstruction through simulations and by comparison with known properties of the diffuse Galactic foregrounds.

The paper is organized as follows. In Sect. 2 we discuss the expected sources of sky emission over the *Planck* frequency interval and how they are modelled. Then in Sect. 3 we detail the overall approach and introduce the four component separation methods. In Sect. 4 we present the *Planck* data set and preprocessing procedure, and we describe our simulations. This is followed by a presentation of the derived CMB maps and their characterization in Sect. 5. Section 6 is dedicated to power spectra and cosmological parameter constraints obtained from these maps, and Sect. 7 to studies of higher-order statistics. Section 8 presents a reconstruction of the diffuse Galactic foregrounds, and Sect. 9 concludes. We relegate details of the algorithms to appendices.

2. The sky at *Planck* frequencies

The properties of Galactic emission vary significantly across the *Planck* frequency range from 30 to 857 GHz. At frequencies below 70 GHz, the dominant radiation processes are: synchrotron emission from cosmic ray electrons interacting with the Galactic magnetic field (e.g., Haslam et al. 1982; Reich & Reich 1988; Broadbent et al. 1989; Davies et al. 1996; Platania et al. 2003; Bennett et al. 2003; Gold et al. 2011); thermal Bremsstrahlung (or free-free emission) from electron-electron and electron-ion scattering (e.g., Banday et al. 2003; Dickinson et al. 2003; Davies et al. 2006; Ghosh et al. 2012; Planck Collaboration Int. XII 2013; Planck Collaboration XX 2011); and AME from dust grains (Kogut 1996; Leitch et al. 1997; Banday et al. 2003; Lagache 2003; de Oliveira-Costa et al. 2004; Finkbeiner et al. 2004; Davies et al. 2006; Bonaldi et al. 2007; Dobler & Finkbeiner 2008; Miville-Deschênes et al. 2008; Ysard et al. 2010; Gold et al. 2011; Planck Collaboration XX 2011), possibly due to their rotational line emission (Draine & Lazarian 1998; Ali-Haïmoud et al. 2009; Ysard & Verstraete 2010; Hoang & Lazarian 2012). Over the frequency range covered by *Planck*, both synchrotron and free-free spectra are well approximated by power laws in brightness temperature, $T_{\rm B} \propto \nu^{\beta}$, with the synchrotron index, $\beta_{\rm synch}$, ranging from -3.2 to -2.8 (Davies et al.

[1] *Planck* (http://www.esa.int/Planck) is a project of the European Space Agency (ESA) with instruments provided by two scientific consortia funded by ESA member states (in particular the lead countries France and Italy), with contributions from NASA (USA) and telescope reflectors provided by a collaboration between ESA and a scientific consortium led and funded by Denmark.

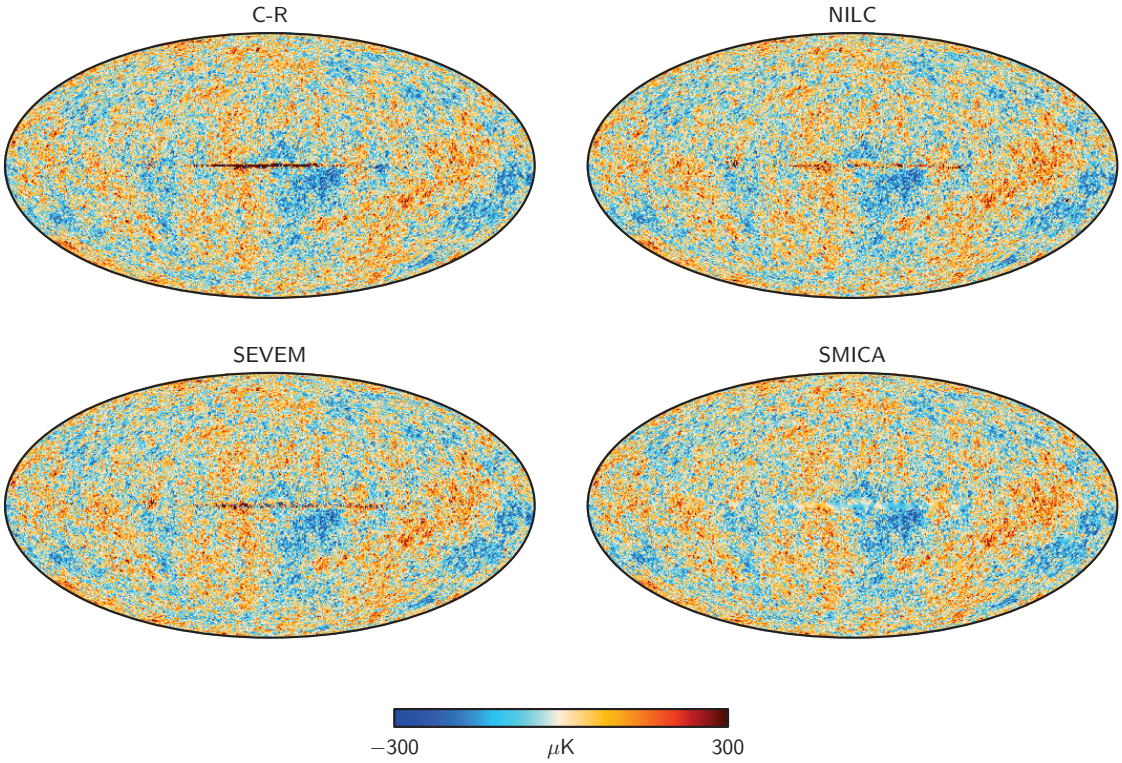


Fig. 1: Foreground-cleaned CMB maps derived by `Commander-Ruler`, `NILC`, `SEVEM` and `SMICA`. Note that the `SMICA` map has been filled in smoothly inside a 3 % Galactic mask.

1996) and the free-free index, $\beta_{\rm ff}$, lying between −2.2 and −2.1. Less is known about the AME spectrum, but spinning dust models with a spectrum peaking at frequencies below 20 GHz (in brightness temperature units) adequately describe current observations[2]. Above the peak, the spectrum appears consistent with a power-law (e.g., Banday et al. 2003; Davies et al. 2006; Dobler & Finkbeiner 2008; Ghosh et al. 2012). In addition to these three, the existence of a fourth low-frequency foreground component, known as the "Galactic haze", has been claimed, possibly due to a hard-spectrum synchrotron population near the Galactic centre (e.g., Finkbeiner 2004; Dobler & Finkbeiner 2008; Pietrobon et al. 2012; Planck Collaboration Int. IX 2013).

At frequencies higher than 100 GHz, thermal dust emission dominates over most of the sky and is commonly described by a modified blackbody spectrum with power-law emissivity, $\epsilon_\nu \propto \nu^{\beta_{\rm d}}$, and temperature, $T_{\rm d}$. Both the temperature and spectral index, $\beta_{\rm d}$, vary spatially. Prior to *Planck*, the best-fitting single component dust model had a temperature $T_{\rm d} \approx 18$ K and spectral index $\beta_{\rm d} \approx 1.7$ (Finkbeiner et al. 1999; Bennett et al. 2003; Gold et al. 2011), although there is evidence of flattening of the spectral index from around 1.8 in the far-infrared to 1.55 in the microwave region (Planck Collaboration 2012), the interpretation of which is still under study.

In addition to these diffuse Galactic components, extragalactic emission contributes at *Planck* frequencies. In particular, a large number of radio and far-infrared (FIR; Planck Collaboration XIII 2011) galaxies, clusters of galaxies and the Cosmic Infrared Background (CIB; Planck Collaboration XVIII 2011) produce a statistically isotropic foreground, with frequency spectra well approximated by models similar to those applicable to the Galactic foregrounds (modified blackbody spectra, power laws, etc.). Except for a frequency-dependent absolute offset, which may be removed as part of the overall offset removal procedure, these extragalactic components are therefore typically absorbed by either the low-frequency or thermal dust components during component separation. No special treatment is given here to extragalactic foregrounds, beyond the masking of bright objects. Dedicated scientific analyses of these sources are described in detail in Planck Collaboration XVIII (2011), Planck Collaboration XXVIII (2013), and Planck Collaboration XXIX (2013). In the *Planck* likelihood, extragalactic sources are modelled in terms of power spectrum templates at high ℓ (Planck Collaboration XV 2013).

Other relevant sources include emission from molecular clouds, supernova remnants, and compact H II regions inside our own Galaxy, as well as the thermal and kinetic SZ effects, due to inverse Compton scattering of CMB photons off free electrons

[2] Note that we adopt brightness temperature for AME in this paper, while many other publications adopt flux density. When comparing peak frequencies, it is useful to note that that a spectrum that has a maximum at 30 GHz in flux density peaks at 17 GHz in brightness temperature.

Table 1: Overview and comparison of component separation algorithms.

Characteristic	Commander-Ruler	NILC	SEVEM	SMICA
Method	Bayesian parameter estimation	Internal linear combination	Internal template fitting	Spectral parameter estimation
Domain	Pixel	Needlet	Pixel	Spherical harmonic
Channels [GHz]	30–353	44–857	30–857	30–857
Effective beam FWHM [arcmin]	~7.4	5.0	5.0	5.0
$\ell_{\rm max}$	none	3200	3100	4000

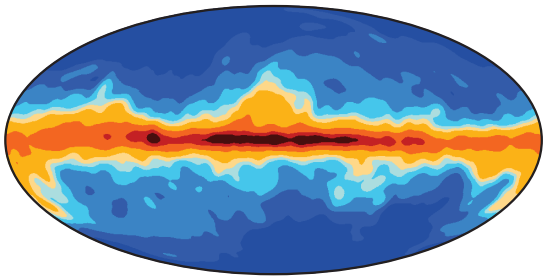


Fig. 2: Combined Galactic (CG) emission masks for the *Planck* data, corresponding to sky fractions of 20, 40, 60, 70, 75, 80, 90, 97, and 99 %. The masks are named CG20, etc.

in ionized media. *Planck* provides new and important information on all these processes, as described both in the following and in the companion papers Planck Collaboration XIII (2013), Planck Collaboration XXI (2013), and Planck Collaboration (2013b). In particular, *Planck*'s frequency range, angular resolution and sensitivity make it a powerful probe of thermal dust, resulting in new and tight constraints on dust temperature and emissivity. The same frequencies also allow extraction of the first ever full-sky maps of the emission resulting from the CO J=1→0, J=2→1 and J=3→2 rotational transitions at 115, 230 and 345 GHz, respectively (Planck Collaboration XIII 2013).

The focus of this paper is to reconstruct the CMB anisotropies over a large sky fraction, exploiting only the *Planck* frequency bands. We also present a detailed reconstruction of the thermal dust emission at high frequencies, as well as CO emission lines. At low frequencies and over the region used for CMB analysis, the total foreground contribution is well approximated by a single power law (see Sect. 8). We therefore model the sum of all low-frequency foregrounds by a power law with spatially varying spectral index whose numerical value in any pixel results from the influence of the dominant foreground component at that location. The full analysis of diffuse foregrounds, using ancillary data to resolve the individual components at low frequencies, will be presented in a forthcoming publication.

3. Approach to component separation

The rich content of the *Planck* data encourages application of several component separation techniques. We consider four, as summarized in Table 1, which we classify according to one of two different general methodologies. The first makes minimal assumptions concerning the foregrounds and seeks only to minimize the variance of the CMB, i.e., the sky component possessing a blackbody spectrum. We implement this approach with a needlet (wavelet on the sphere) version of the internal linear combination (ILC) algorithm (NILC; Delabrouille et al. 2009), and also with a template-based method to remove foreground contamination from the CMB-dominant bands. These foreground templates are constructed from the lowest and highest frequency channels (Spectral Estimation Via Expectation Maximization, SEVEM; Fernández-Cobos et al. 2012).

The second methodology uses parametric modelling of the foregrounds. In our real space implementation, we explore model parameters through Bayesian parameter estimation techniques, fitting a parametric signal model per pixel (Commander; Eriksen et al. 2006, 2008); a similar implementation is presented by Stompor et al. (2009). To estimate spectral indices robustly in pixel space, this procedure requires identical angular resolution across all frequencies included in the analysis, and is therefore limited in resolution by the 30 GHz LFI channel. However, this is sufficient to generate the low-resolution CMB map and power spectrum samples required for the low multipole part of the *Planck* likelihood function for cosmological parameters (Planck Collaboration XV 2013). To produce full resolution maps, we use the resulting low-resolution spectral parameter samples to solve for the component amplitudes, in an extension to the method known as Ruler (we refer to the combined method as Commander-Ruler, or C-R). In our fourth technique, we implement a CMB-oriented parametric approach that fits the amplitude and spectral parameters of CMB and foregrounds in the harmonic domain (Spectral Matching Independent Component Analysis, SMICA; Cardoso et al. 2008).

Details of each algorithm are given in the appendices. We now turn to their application to the data and evaluate their performance using simulations.

4. Data, simulations and masks

We use the data set from the first 15.5 months of *Planck* observations, corresponding to 2.6 sky surveys, from both the Low Frequency Instrument (LFI) and High Frequency Instrument (HFI). The primary inputs for component separation are the frequency channel maps, including half-ring maps, bandpasses, and beam characteristics; a full description of these products is given in Planck Collaboration II (2013) and Planck Collaboration VI (2013). No special corrections are made for zodiacal light emission (ZLE; Planck Collaboration VI 2013) in the analyses presented here. The ZLE is not stationary on the sky, since it depends on *Planck*'s position and scanning strategy. Therefore the frequency maps contain a projected version of the emission averaged over the nominal mission. Despite this, a series of exploratory analyses showed that our algorithms naturally correct for this component within their existing model space. It was also found that larger CMB residuals were induced when applying a correction based on a ZLE model than when applying no correction, most likely due to uncertainties in the model itself.

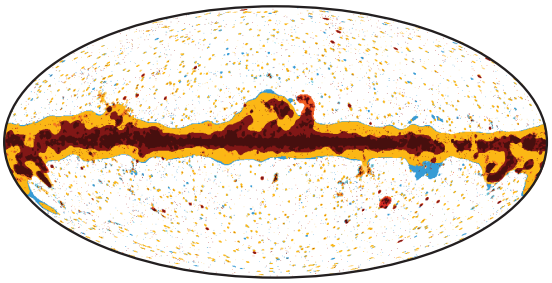


Fig. 3: Summary of Component Separation (CS) confidence masks. Each pixel is encoded in terms of a sum in which `Commander-Ruler` equals 1 (light blue), `NILC` equals 2 (dark red), `SEVEM` equals 4 (yellow), and `SMICA` equals 8 (light red). The masks are named CS-CR75, CS-NILC93, CS-SEVEM76, and CS-SMICA89, respectively, reflecting their accepted sky fraction. The union mask (U73), used for evaluation purposes in this paper, removes all coloured pixels.

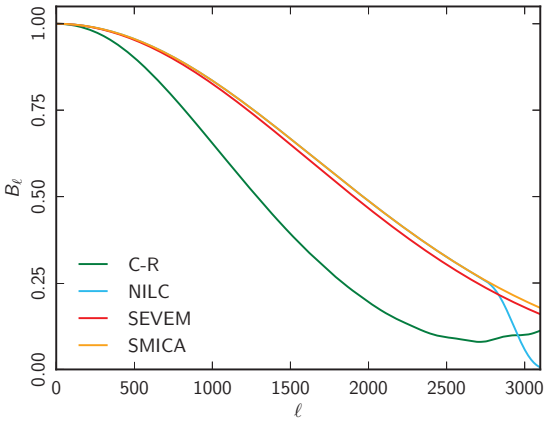


Fig. 4: Beam transfer functions of the four foreground-cleaned CMB maps.

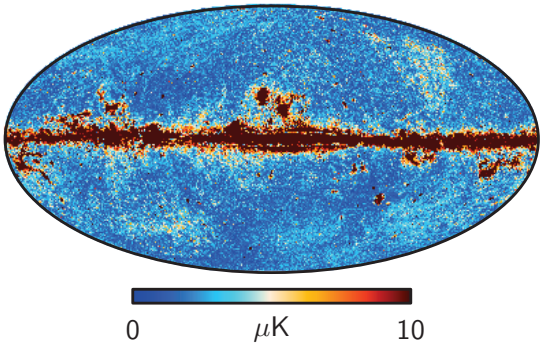


Fig. 5: Standard deviation between the four foreground-cleaned CMB maps. All maps have been downgraded to a `HEALPix` resolution of $N_{\rm side} = 128$. The differences are typically less than 5 μK at high Galactic latitudes, demonstrating that the maps are consistent over a large part of the sky.

To evaluate and validate our algorithms, we analyse a large suite of realistic simulations, the so-called Full Focal Plane (FFP) simulations, based on detailed models of the instrument and sky. The version used for this data release is denoted FFP6, and is described in Planck Collaboration ES (2013). The simulation procedure generates time streams for each detector, incorporating the satellite pointing, the individual detector beams, bandpasses, noise properties, and data flags, and then produces simulated frequency channel maps through the mapmaking process. For the input sky, we use the *Planck* Sky Model (PSM), which includes the CMB, diffuse Galactic emission (synchrotron, free-free, thermal dust, AME, and molecular CO lines), and compact sources (thermal and kinetic SZ effects, radio sources, infrared sources, the CIB, and ultra-compact H II regions). The pre-launch version of the PSM is described by Delabrouille et al. (2012), and has been modified for the present work as described in Planck Collaboration ES (2013). Each FFP data set consists of three parts: the simulated observations, Monte Carlo realizations of the CMB, and Monte Carlo realizations of the instrumental noise.

For both the data and the simulations, we reconstruct the CMB and foregrounds from the full frequency channel maps and the corresponding half-ring maps, which are made from the data in the first half or second half of each stable pointing period. The half-ring maps can be used to obtain an estimate of the noise in each channel by taking half of the difference between the two maps, thereby normalizing the noise level to that of the full map. This is referred to as the half-ring half-difference (HRHD) map. The signals fixed to the sky will be cancelled leaving only the noise contribution. The HRHD map can be treated as a realization of the same underlying noise processes and it can be used to estimate the power spectrum, and other properties, of the noise. If there are noise correlations between the half-ring maps, then the estimates of the noise properties thus obtained can be biased. This is the case for HFI channels; the cosmic ray glitch removal (Planck Collaboration VI 2013; Planck Collaboration X 2013) induces correlations that lead to the noise power spectrum being underestimated by a few percent at high ℓ when using the HRHD maps.

Prior to processing the data through each component separation pipeline, we define masks for the point sources and bright Galactic regions. Point source masking is based on the source catalogues obtained by filtering the input sky maps with the Mexican Hat Wavelet 2 (MHW2) filter and applying a 4σ threshold for the LFI bands and a 5σ threshold for the HFI bands (Planck Collaboration XVIII 2011; Planck Collaboration XXVIII 2013). The mask radius of each source is different for the LFI and HFI. Due to the large beam size of LFI channels, we define a variable masking radius for each source according to its signal-to-noise ratio (S/N) as $r = (2\log(A/m))^{1/2}/(2\sqrt{2\log 2})\times\mathrm{FWHM}$, where r is the radius, A is the S/N, and m is the maximum amplitude (given in units of the background noise level) allowed for the tail of unmasked point sources; we set $m = 0.1$, which is a compromise between masking the source tails and minimizing the number of masked pixels.

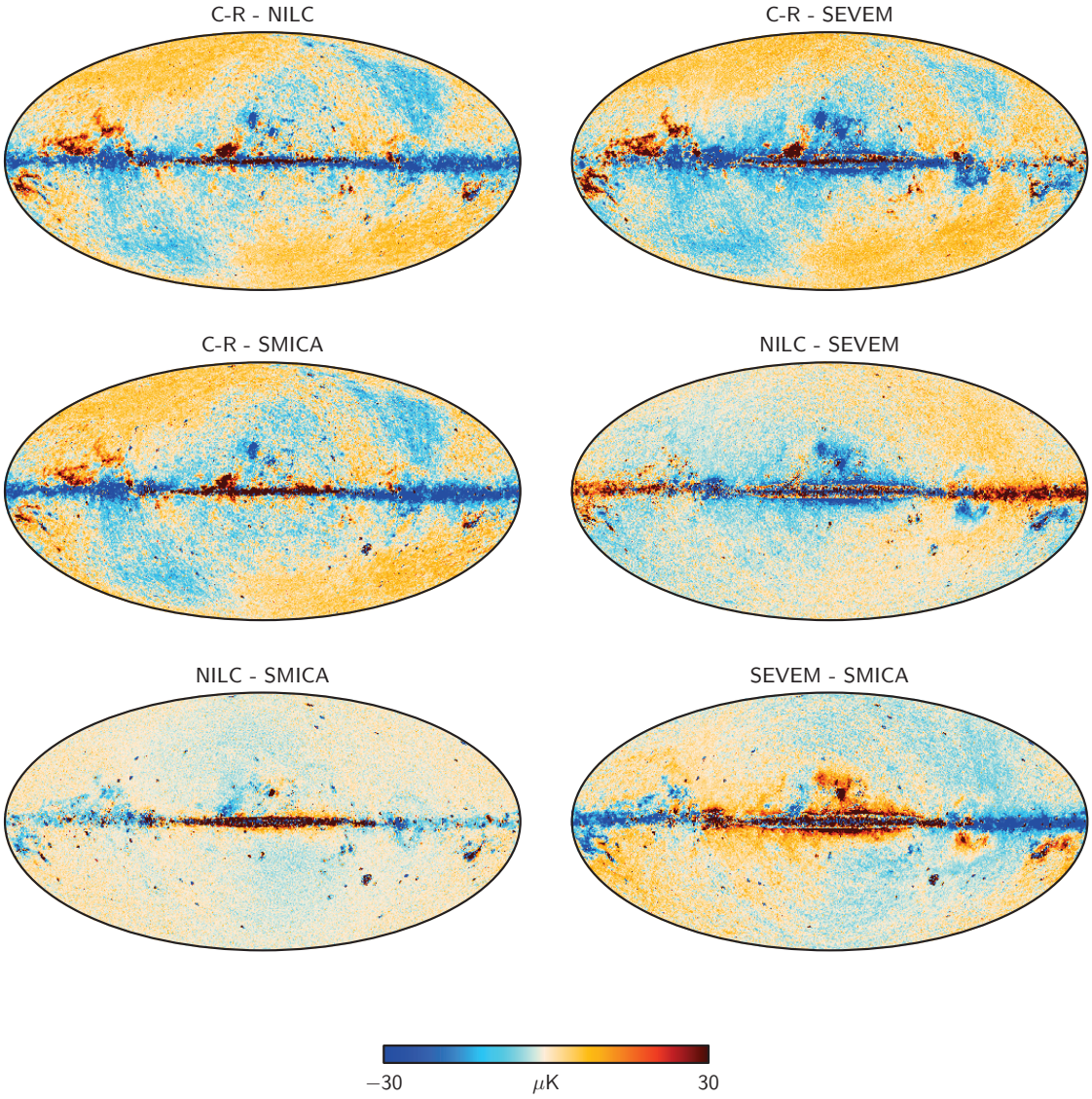


Fig. 6: Pairwise differences between foreground-cleaned CMB maps. All maps have been downgraded to a `HEALPix` resolution of $N_{\rm side} = 128$ to show the large-scale differences. The line-like discontinuities in the differences involving `SEVEM` is due to the two different regions used in this algorithm to clean the sky (see Appendix C for details).

For HFI, the mask radius around each source is 1.27×FWHM, using the average FWHM obtained from the effective beams.

A basic set of Galactic masks is defined as follows. We subtract a CMB estimate from the 30 and 353 GHz maps, mask point sources, and smooth the resulting maps by a Gaussian with FWHM of 5°. We then threshold and combine them, generating a series of masks with different amounts of available sky. The resulting combined Galactic (CG) masks, shown in Fig. 2, correspond to sky fractions of 20, 40, 60, 70, 75, 80, 90, 97, and 99 %, and are named CG20, etc.

5. CMB Maps

We begin the discussion of our results by presenting the foreground-cleaned CMB maps. These maps are shown in Fig. 1 for each of the four component separation algorithms. Already from this figure it is clear that the wide frequency coverage and high angular resolution of *Planck* allow a faithful reconstruction

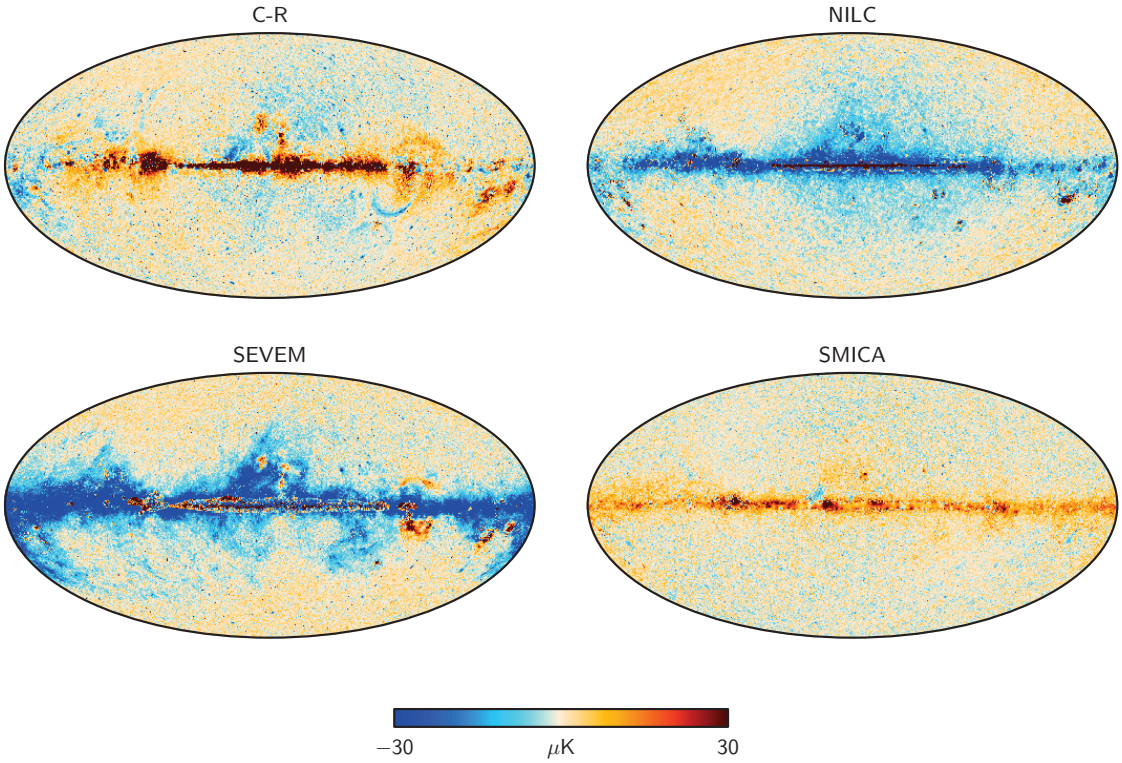


Fig. 7: CMB residual maps from the FFP6 simulation. A monopole determined at high Galactic latitude has been subtracted from the maps, and they have been downgraded to a HEALPix resolution of $N_{\rm side} = 128$ to show the large-scale features. The residuals presented here provide a conservative estimate of those expected in the data (see text for details).

of the CMB field over most of the sky. The fluctuations appear visually consistent with the theoretical expectation of a Gaussian and isotropic signal everywhere except inside a small band very close to the Galactic plane[3]

Each CMB map is accompanied by its own confidence mask outside which the corresponding solution is considered statistically robust, shown in Fig. 3; for a definition of each mask, see Appendices A–D. Accepted sky fractions are 75, 93, 76, and 89 %, respectively, for Commander-Ruler, NILC, SEVEM, and SMICA. These masks are denoted CS-CR75, CS-NILC93, CS-SEVEM76, CS-SMICA89, respectively. The union of the confidence masks accepts 73 % of the sky and is denoted U73. It is adopted as the default mask for evaluation purposes in this paper.

In addition to the CMB maps from the full data set, the half-ring frequency maps have been processed by each algorithm to provide half-ring CMB maps. They are used to provide estimates of the instrumental noise contribution to the foreground-cleaned maps in the power spectrum analysis (see Sect. 6). The algorithms were also used to process Monte Carlo simulations: 1000 realizations of the CMB and 1000 realizations of noise. They are not used in the analyses presented in this paper, but are used by Planck Collaboration XXIII (2013) and Planck Collaboration XXIV (2013).

The beam transfer functions of the foreground-cleaned CMB maps have been estimated for each algorithm, as shown in Fig. 4. The angular resolution of the NILC, SEVEM, and SMICA maps corresponds to a Gaussian beam with FWHM of 5′. The difference between SEVEM and NILC/SMICA is due to their different treatment of the HEALPix[4] pixel window function (Górski et al. 2005). The deviation of NILC beam from a Gaussian shape at $\ell > 2800$ is caused by the last needlet window (see Appendix B). Commander-Ruler has a larger beam, because it is defined explicitly as a weighted average of frequency maps in pixel space. Its resolution is equivalent to a Gaussian beam with FWHM of approximately 7.′4. The beam transfer functions have been computed assuming the best-fit beam transfer function for each frequency channel, and the uncertainties in the latter have not been propagated to these estimates.

In Fig. 5 we show the standard deviation per pixel among the four foreground-cleaned CMB maps downgraded to $N_{\rm side} = 128$, and in Fig. 6 we show all pairwise difference maps. Typical differences at high Galactic latitudes are smaller than 5 μK. Considering the difference maps in more detail, it is clear that

[3] Note that SMICA, being defined in harmonic space, employs a smooth filling process inside a small Galactic mask to prevent foreground residuals from leaking from low to high Galactic latitudes, and therefore appears visually different from the other three solutions in this respect; see Appendix D.

[4] http://healpix.sourceforge.net

the Commander-Ruler map is the most different from the other three, whereas NILC and SMICA are the most similar. This is not completely unexpected, because while Commander-Ruler uses only frequencies between 30 and 353 GHz in its solution, the other three codes additionally include the dust-dominated 545 and 857 GHz maps.

This difference in data selection may explain some of the coherent structures seen in Fig. 6. In particular, the most striking large-scale feature in the difference maps involving Commander-Ruler is a large negative band roughly following the ecliptic plane. This is where the ZLE (Planck Collaboration VI 2013) is brightest. Since the ZLE is also stronger at high frequencies, having a spectrum close to that of thermal dust, it is possible that this pattern may be an imprint of residual ZLE either in the Commander-Ruler map, or in all of the other three maps. Both cases are plausible. The Commander-Ruler solution may not have enough high-frequency information to distinguish between ZLE and normal thermal dust emission, and, by assuming a thermal dust spectrum for the entire high-frequency signal at 353 GHz, over-subtracts the ZLE at lower frequencies. It is also possible that the other three CMB solutions have positive ZLE residuals from extrapolating the high-frequency signal model from 857 GHz to the CMB frequencies. Without an accurate and detailed ZLE model, it is difficult to distinguish between these two possibilities. It is of course also possible that the true explanation is in fact unrelated to ZLE, and the correlation with the ecliptic plane is accidental. In either case, it is clear that the residuals are small in amplitude, with peak-to-peak values typically smaller than $10\,\mu$K, of which by far the most is contained in a quadrupole aligned with the ecliptic. This provides additional evidence that residual ZLE is not important for the CMB power spectrum and cosmological parameter estimation, although some care is warranted when using these maps to study the statistics of the very largest angular scales (e.g., Planck Collaboration XXIII 2013); checking consistency among all four maps for a given application alleviates much of this concern.

We end this section by showing in Fig. 7 a set of residual maps derived by analysing the FFP6 simulation with exactly the same analysis approaches as applied to the data. It is evident that SMICA produces the map with lowest level of residuals. Considering the morphology in each case, we see that the main contaminant for Commander-Ruler is under-subtracted free-free emission, while for both NILC and SEVEM it is over-subtracted thermal dust emission, and for SMICA it is under-subtracted thermal dust emission. However, at high latitudes and outside the confidence masks, the residuals are generally below a few μK in amplitude. It is also worth noting that each algorithm has been optimized (in terms of model definition, localization parameters, etc.) for the data, and the same configuration was subsequently used for the FFP6 simulations without further tuning. The simulations presented here therefore provide a conservative estimate of the residuals in the data. This is also reflected in the fact that the differences between CMB reconstructions for the FFP6 simulations are larger than those found in the data. See Appendix E for further details.

6. Power spectrum and cosmological parameters

In this section we evaluate the foreground-cleaned maps in terms of CMB power spectra and cosmological parameters. Our purpose in doing this is to show that the maps are consistent with the high-ℓ likelihood obtained from the cross-spectrum analysis of detector set and frequency maps in Planck Collaboration XV (2013), and with the cosmological parameters derived from them in Planck Collaboration XVI (2013). This also establishes the consistency between *Planck*'s cosmological constraints and studies of the large-scale structure and higher order statistics of the CMB.

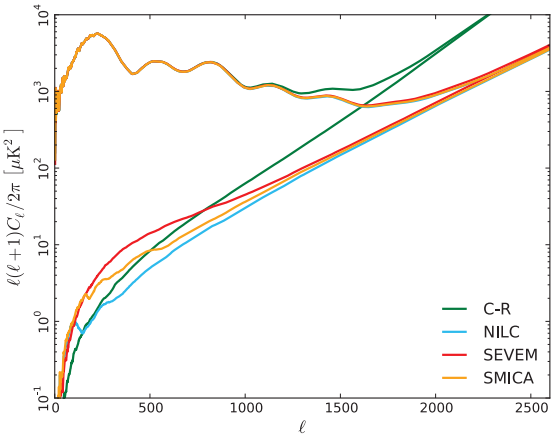


Fig. 8: Angular power spectra of the foreground-cleaned CMB maps and half-ring half-difference (HRHD) maps. The spectra have been evaluated using the U73 mask apodized with a 30′ cosine function.

6.1. Power spectra

Figure 8 shows the power spectra of the foreground-cleaned CMB maps and the corresponding HRHD maps, evaluated using the U73 mask with a 30′ cosine apodization. The spectra have been corrected for the effect of the mask and the beam transfer function of each algorithm has been deconvolved. The spectra of the HRHD maps give an estimate of the instrumental noise contribution to the power spectrum of the cleaned map. The correlations between the HFI half-ring frequency maps are inherited by the half-ring CMB maps that use them as input. At small angular scales, the CMB solution comes almost entirely from data in the HFI channels, and therefore the spectrum of the CMB HRHD maps is also biased low.

At small angular scales, the effective noise levels of NILC, SEVEM, and SMICA are very similar, and lower than that of Commander-Ruler. The last has larger noise because it operates entirely in pixel space and therefore applies the same weights to all multipoles. It cannot take advantage of the changing signal-to-noise ratio of the frequency channels with angular scale.

We can estimate the contribution of residual foregrounds to the foreground-cleaned CMB maps by making use of the FFP6 simulations. In addition to processing the simulated frequency maps, the maps of the individual input sky components were processed by the algorithms after fixing their parameters or weights to the values obtained from the "observed" maps. Figure 9 shows the power spectra of the simulated FFP6 components, in this case CMB, noise and the sum of the foreground components. The top panel shows the spectra computed using the union mask derived from the simulation with a 30′ cosine apodization. The total foreground contribution becomes comparable to the CMB signal at $\ell \approx 2000$. The bottom panel shows the same computed with an apodized point source mask applied to the maps (i.e.,

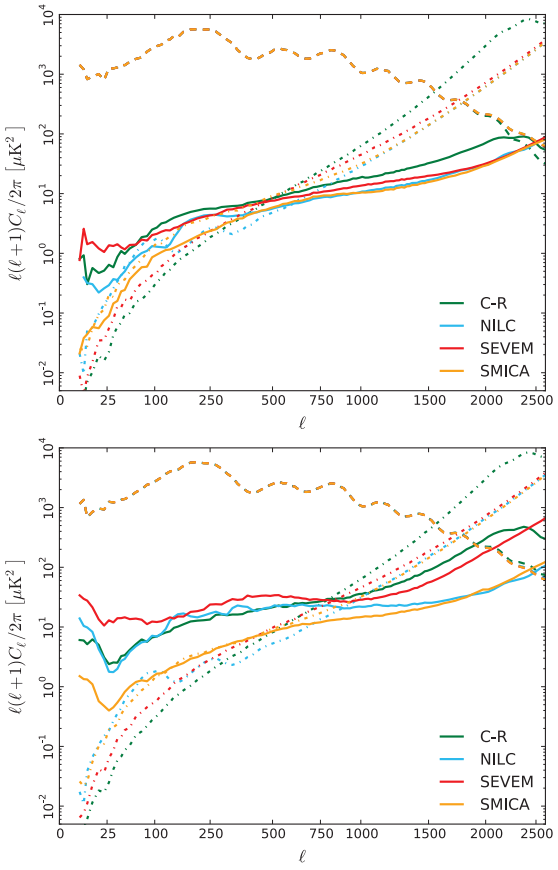


Fig. 9: Angular power spectra of FFP6 simulated components evaluated over the common mask (top) and the common point source mask (bottom), both apodized with a 30′ cosine function. Three components are shown: the CMB (dashed line); noise (dot-dashed line); and the sum of all foregrounds (solid line). A nonlinear scale is used on the horizontal axis to show all the features of the spectra.

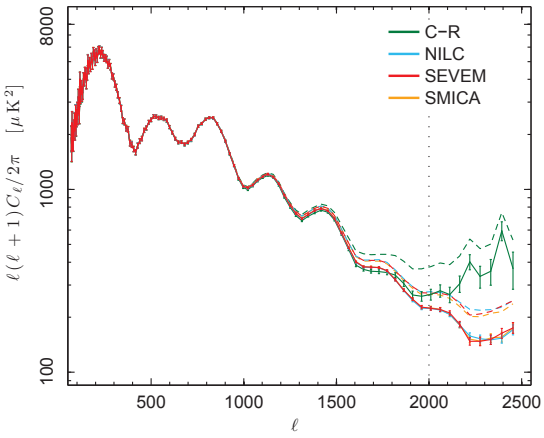


Fig. 10: Estimates of the CMB power spectra from the foreground-cleaned maps, computed by `XFaster`. The solid lines show the spectra after subtracting the best-fit model of residual foregrounds. The vertical dotted line shows the maximum multipole ($\ell = 2000$) used in the likelihood for fitting the foreground model and cosmological parameters (see Sect. 6.2.2 for further details). The dashed lines show the spectra before residual foreground subtraction.

no diffuse masking, although this mask does removes a large part of the Galactic plane). The residual foreground contribution is larger at all angular scales, but still it only becomes comparable to the CMB signal at $\ell \approx 1800$ in the worst case. For both masks, `SMICA` has the smallest residual foreground contamination at large angular scales, which is also demonstrated in Fig. 7. A more detailed examination of the contribution of the individual foreground components to the power spectrum is in Appendix E.

6.2. Likelihood and cosmological parameters

We estimate the binned power spectra with `XFaster` (Rocha et al. 2009, 2010, 2011) and determine cosmological parameter constraints using a correlated Gaussian likelihood. Parameter constraints are derived using a Metropolis-Hastings Markov Chain Monte Carlo sampler. To speed up this process, we additionally use `PICO` (Parameters for the Impatient COsmologist, Fendt & Wandelt 2008), a tool which interpolates the CMB power spectra and matter power spectra as a function of cosmological parameters.

6.2.1. Model and methods

We compute the power spectrum for each foreground-cleaned map over the multipole range $2 \le \ell \le 2500$, while parameter constraints are derived using only $70 \le \ell \le 2000$; as shown in Appendix E through simulations, modelling errors become non-negligible between $\ell = 2000$ and 2500. For parameter estimation, we adopt a standard six-parameter ΛCDM model, and impose an informative Gaussian prior of $\tau = 0.0851 \pm 0.014$, since polarization data are not included in this analysis.

While the foreground-cleaned maps should have minimal contamination from diffuse Galactic emission, they do contain significant contamination from unresolved extragalactic sources. These contributions are most easily modelled in terms of residual power spectra, therefore we marginalize over the corresponding parameters at the power spectrum level. To the six ΛCDM parameters, describing the standard cosmology, we add two foreground parameters, $A_{\rm ps}$, the amplitude of a Poisson component (and hence constant, $C_\ell = A_{\rm ps}$), and $A_{\rm cl}$, the amplitude of a clustered component with shape $D_\ell = \ell(\ell+1)C_\ell/2\pi \propto \ell^{0.8}$. Both are expressed in terms of D_ℓ at $\ell = 3000$ in units of μK^2.

The power spectrum calculation is based on the half-ring half-sum (HRHS) and HRHD CMB maps (see Sect. 5); the latter is used to estimate the noise bias in the power spectra extracted from the HRHS maps. From these, we calculate the pseudo-spectra, $\tilde{C}_\ell$ and $\tilde{N}_\ell$ (Hivon et al. 2002), respectively, after applying the U73 mask. These are used as inputs to `XFaster` together

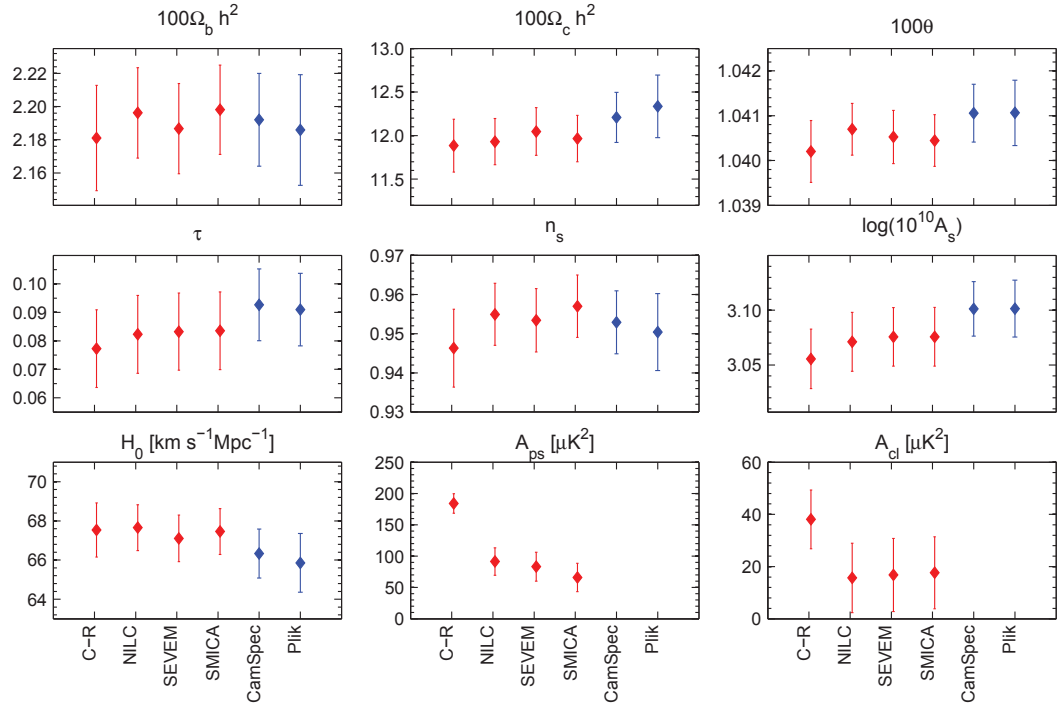


Fig. 11: Comparison of cosmological and foreground parameter values estimated from the foreground-cleaned CMB maps for $\ell_{\rm max} = 2000$ (in red) and those obtained with `CamSpec` and `Plik` likelihoods (in blue). The values of the foreground parameters are not shown for `CamSpec` and `Plik`, since they use a different foreground model.

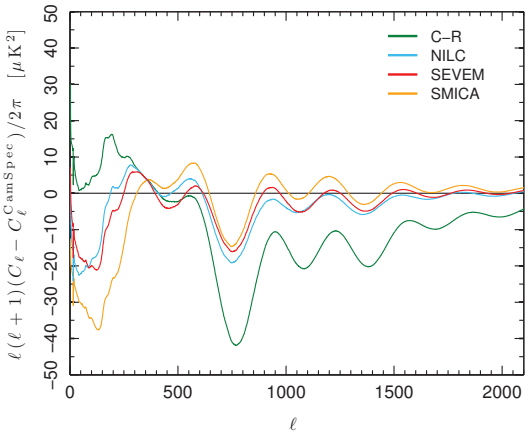


Fig. 12: Residuals of all map-based best-fit models relative to `CamSpec` best-fit model (assuming a prior on τ) for $\ell_{\rm max} = 2000$.

with the beam transfer functions provided by each method (see Fig. 4).

To avoid aliasing of power from large to small scales, which would add an offset between the signal-plus-noise and noise pseudo-spectra at high ℓ, we use the apodized version of the U73 mask. The known mismatch in the noise level between the spectra due to the correlation between the half-ring maps is not explicitly corrected. It is left to be absorbed into the two foreground parameters.

Using the pseudo-spectra and `XFaster`, we then reconstruct an estimate of the power spectrum of each foreground-cleaned HRHS map, removing the noise bias as estimated from the corresponding HRHD map. To this end we apply an iterative scheme starting from a flat spectrum model. The result is a binned power spectrum and the associated Fisher matrix, which are then used to construct the likelihood, approximated here by a correlated Gaussian distribution.

To study consistency in the low-ℓ range, we fit a two-parameter q–n (amplitude-tilt) model relative to the *Planck* best-fit ΛCDM model on the form, $C_\ell = q(\ell/\ell_{\rm pivot})^n C_\ell^{\rm bf}$, using a pixel-space likelihood for maps smoothed to 6° FWHM; see Planck Collaboration XV (2013) for further algorithmic details.

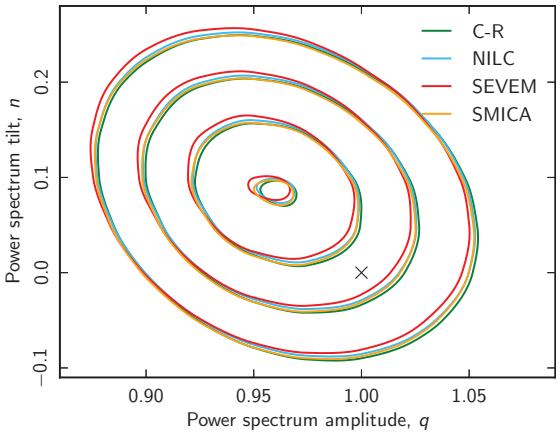


Fig. 13: Low-ℓ power spectrum amplitude and tilt constraints measured relative to the best-fit *Planck* ΛCDM model derived from foreground-cleaned CMB maps smoothed to 6° FWHM. The cross shows the best-fit model $(q, n) = (1, 0)$.

6.2.2. Results

We perform the power spectrum and parameter estimation analysis for both the data and the FFP6 simulations described in Sect. 4. The results for the latter are given in Appendix E.

Figure 10 shows estimates of the angular power spectrum for each foreground-cleaned map, with the uncertainties given by the Fisher matrix. The parameter summary given in Fig. 11 shows the parameter constraints derived using multipoles between $\ell = 70$ and 2000, and compares these to results obtained with the `CamSpec` and `Plik` likelihoods (Planck Collaboration XV 2013).

Differences in the power spectra at high ℓ are mostly absorbed by the two-parameter foreground model, rendering consistent cosmological parameters. For example, the high-ℓ power excess seen in the `Commander-Ruler` map is well-fitted in terms of residual point sources, which makes intuitive sense, considering the lower angular resolution of this map (see Sect. 5). However, the ΛCDM parameter uncertainties derived from the four codes are very consistent. This indicates that most of the cosmological information content above $\ell \geq 1500$ is degenerate with the extragalactic foreground model, and a more sophisticated foreground treatment is required in order to recover significant cosmological parameter constraints from these scales. Beyond this, deviations among cosmological parameters are small and within $1\,\sigma$ for all methods and most of the parameters. Further, the parameters derived from the four foreground-cleaned CMB maps are in good agreement with those obtained by `CamSpec` and `Plik` using cross-spectra; departures are well within $1\,\sigma$ for most parameters.

Inspecting the differences between the best-fit models derived from the four foreground-cleaned maps and from `CamSpec` plotted in Fig. 12, we find that the relative residuals are within $40\,\mu\mathrm{K}^2$ for all multipole ranges, and smaller than $20\,\mu\mathrm{K}^2$ at high ℓ. This can be compared to the corresponding residuals for the FFP6 simulation shown in Appendix E.

The likelihood used for this analysis does not take into account some systematic effects that will affect our foreground-cleaned CMB maps, such as relative calibration uncertainties between the frequency channel maps used to construct them, or their beam uncertainties. These effects are accounted for in the likelihoods in Planck Collaboration XV (2013). We have also adopted a very simple two-parameter model for the residual extragalactic foregrounds. Despite these limitations, the four CMB maps yield cosmological parameters in agreement with the cross-spectrum based likelihoods for a basic six-parameter ΛCDM model. Thus we can be confident that the CMB maps are consistent with the power spectrum analysis.

Before concluding this section, we show in Fig. 13 the results from a two-parameter fit of an amplitude-tilt model to each of the four foreground-cleaned maps, downgraded to 6° and repixelized at an $N_{\mathrm{side}} = 32$ grid. Clearly, the maps are virtually identical on large angular scales measured relative to cosmic variance, with any differences being smaller than $0.1\,\sigma$ in terms of cosmological parameters. However, it is worth noting that the best-fit model, $(q, n) = (1, 0)$, is in some tension with the low-ℓ spectrum, at about $1.7\,\sigma$ in this plot. The same tension between large and small angular scales is observed in Planck Collaboration XV (2013) and Planck Collaboration XVI (2013) with higher statistical significance using the full *Planck* likelihood. Irrespective of physical interpretation, the calculations presented here demonstrate that these low-ℓ features are robust with respect to component separation techniques.

7. Higher-order statistics

The foreground-cleaned CMB maps presented in this paper are used as inputs for most *Planck* analyses of higher-order statistics, including non-Gaussianity studies (Planck Collaboration XXIV 2013), studies of statistical isotropy (Planck Collaboration XXIII 2013), gravitational lensing by large-scale structure (Planck Collaboration XVII 2013), and of the integrated Sachs-Wolfe effect (Planck Collaboration XXIV 2013). In this section we provide a summary of the non-Gaussianity and gravitational lensing results.

7.1. Non-Gaussianity

Primordial non-Gaussianity is typically constrained in terms of the amplitude, $f_{\mathrm{NL}}^{\mathrm{local}}$, of the quadratic corrections to the gravitational potential, as well as by means of the three-point correlation function based on different triangle configurations. The results from these calculations for the foreground-cleaned CMB maps are presented in Planck Collaboration XXIV (2013). After subtraction of the lensing-ISW correlation contribution, the final result is $f_{\mathrm{NL}}^{\mathrm{local}} = 2.7 \pm 5.8$, as estimated from the `SMICA` map using the KSW bispectrum estimator (Komatsu et al. 2005), consistent within $1\,\sigma$ with results from other methods and foreground-cleaned maps.

Uncertainties are evaluated by means of the FFP6 simulations, and potential biases are studied using both Gaussian and non-Gaussian CMB realizations. In particular, when a detectable level of primordial non-Gaussianity ($f_{\mathrm{NL}}^{\mathrm{local}} = 20.4075$) is injected into the FFP6 simulations, each foreground-cleaned map yielded a positive detection within $2\,\sigma$ of the expected value, recovering values of $f_{\mathrm{NL}}^{\mathrm{local}} = 8.8 \pm 8.6$, 19.0 ± 7.5, 11.1 ± 7.6 and 19.7 ± 7.4 for `Commander-Ruler`, `NILC`, `SEVEM`, `SMICA`, respectively. We see that `NILC` and `SMICA` demonstrate the best recovery of the injected non-Gaussianity, and we favoured the latter for non-Gaussian studies for its faster performance over `NILC`. The foreground-cleaned CMB maps presented in this paper do not provide significant evidence of a non-zero value of

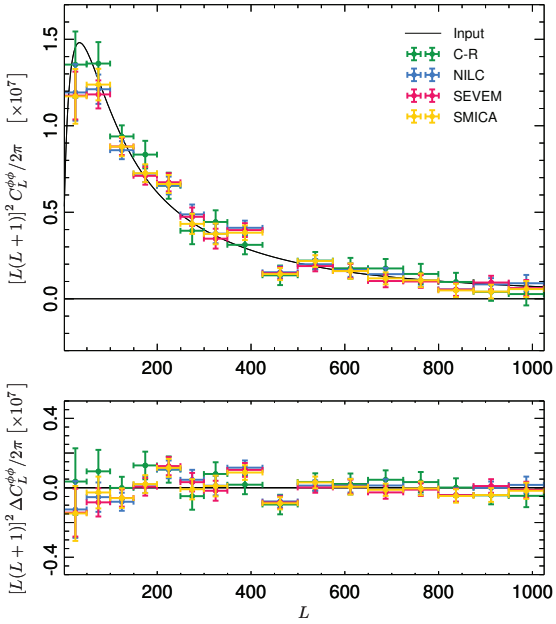


Fig. 14: Lensing power spectrum estimates from FFP6 simulations using an apodized mask covering $f_{\rm sky,2} \simeq 0.70$ of the sky.

$f_{\rm NL}^{\rm local}$, and realistic simulations show that the component separation methods do not suppress real non-Gaussian signatures within expected uncertainties. The implications of these results in terms of early Universe physics are discussed in the relevant papers (Planck Collaboration XXIV 2013; Planck Collaboration XXII 2013).

7.2. Gravitational lensing by large-scale structure

Gravitational lensing by the intervening matter imprints a non-Gaussian signature in the CMB, which allows the reconstruction of the gravitational potential integrated along the line of sight to the last scattering surface. In Planck Collaboration XVII (2013), this effect has been detected at a high significance level (greater than $25\,\sigma$) using the *Planck* temperature maps. Specifically, the lensing induced correlations between the total intensity and its gradients have been used to reconstruct a nearly full sky map of the lensing potential ϕ, which has been used for further studies on *Planck* data, including the detection of a non-zero correlation with the ISW (Planck Collaboration XXIV 2013; Planck Collaboration XIX 2013) and other tracers of large-scale structure (notably, significant correlation with the CIB is reported in Planck Collaboration XVIII 2013), as well as the estimate of the power spectrum of the lensing potential and the associated likelihood. The latter was constructed using a simple minimum variance combination of the 143 and 217 GHz maps on about 70 % of the sky, as well as subtracting dust contamination using the 857 GHz *Planck* channel as a template (Planck Collaboration XVII 2013). These lensing results have improved the cosmological constraints from *Planck* (Planck Collaboration XVI 2013).

The foreground-cleaned CMB maps described in Sect. 5 were used to perform a lensing extraction on a larger sky fraction, reaching about 87 % of the sky. We found the lensing power spectrum to be in good agreement with the one obtained using the minimum variance combination, i.e., the signal agrees within $1\,\sigma$ in the majority of the angular domain bins, and is characterized by an equivalent uncertainty. The foreground-cleaned maps were further exploited on the baseline 70 % sky fraction for assessing the robustness of the main reconstruction against the foreground contamination (Planck Collaboration XVII 2013).

We show that the component separation algorithms presented in this paper do not bias the lensing reconstruction in the case of the large sky fraction considered here. We consider FFP6 simulations including noise and lensed CMB signal, propagated through each of the component separation algorithms described in Sect. 3. We perform a lensing potential reconstruction in the pixel domain based on the CMB maps processed by the four component separation methods using the `metis` algorithm described in Planck Collaboration XVII (2013). This method uses the quadratic estimator presented in Okamoto & Hu (2003), which corrects for the mean-field bias caused by extra sources of statistical anisotropy in addition to the CMB.

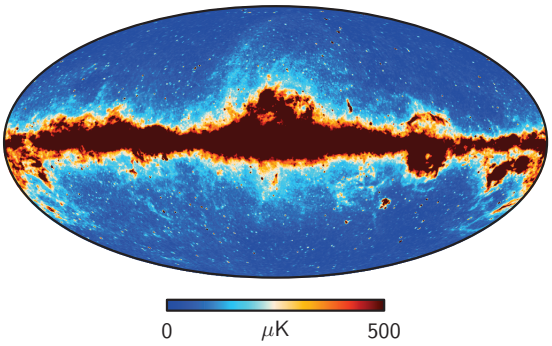
For each method, we combine the masks of CO regions, nearby galaxies and compact objects as defined in Planck Collaboration XVII (2013), with the CG90 mask described in Sect. 4. This procedure results in masks with sky fractions $f_{\rm sky} = 0.836, 0.851, 0.850, 0.846$ for `Commander-Ruler`, `NILC` `SEVEM`, and `SMICA`, respectively.

We estimate the lensing potential power spectrum, $C_L^{\phi\phi}$, following the methodology described in Planck Collaboration XVII (2013). It consists of a pseudo-C_ℓ estimate based on a highly-apodized version of the lensing potential reconstruction, which has an effective available sky fraction $f_{\rm sky,2} = 0.648, 0.690, 0.686, 0.683$ for `Commander-Ruler`, `NILC`, `SEVEM` and `SMICA`, respectively. The band-power reconstructions in 17 bins in the range $2 \le \ell \le 1025$ are plotted in Fig. 14, as well as the residuals relative to the theoretical lens power spectrum. All algorithms achieved an unbiased estimation of the underlying lensing power spectrum, with $\chi^2 = 10.58, 17.34, 18.54, 15.30$, for `Commander-Ruler`, `NILC`, `SEVEM`, and `SMICA` respectively, with 17 degrees of freedom. The associated probability-to-exceed (PTE) values are 83 %, 36 %, 29 %, 50 %.

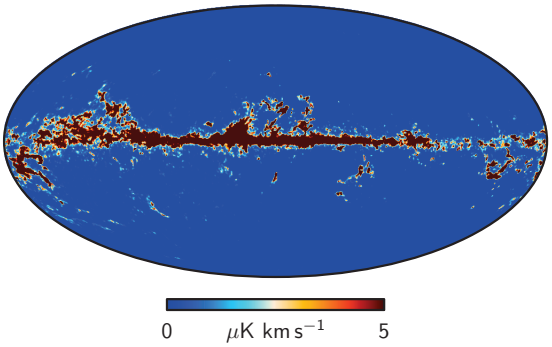
The power spectrum estimates are in remarkable agreement with each other. However, the `Commander-Ruler` solution has significantly larger uncertainties, as expected from its lower signal-to-noise ratio to lensing due to its larger beam. These results on simulated foreground-cleaned CMB maps demonstrate that the component separation algorithms do not alter the lensing signal, and this provides a strategy for achieving a robust lensing reconstruction on the largest possible sky coverage. The foreground-cleaned maps have been used in Planck Collaboration XVII (2013) to obtain lensing potential estimates on 87 % of the sky.

8. Foreground components

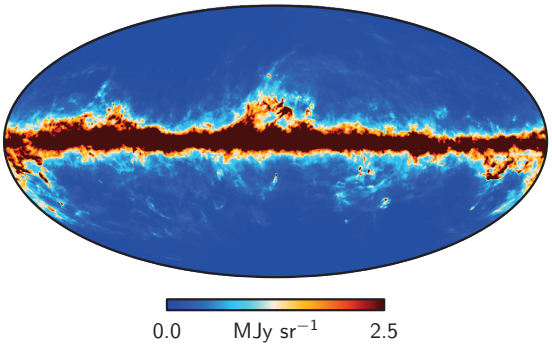
In this section we consider the diffuse Galactic components, and present full-sky maps of thermal dust and CO emission, as well as a single low-frequency component map representing the sum of synchrotron, AME, and free-free emission. Our all-sky CO map is a "type 3" product as presented in Planck Collaboration XIII (2013). To assess the accuracy of these maps, we once again take advantage of the FFP6 simulation. The `Commander-Ruler` method used in the following is described in Appendix A and consists of a standard parametric Bayesian MCMC analysis at



(a) Low-frequency component amplitude at 30 GHz

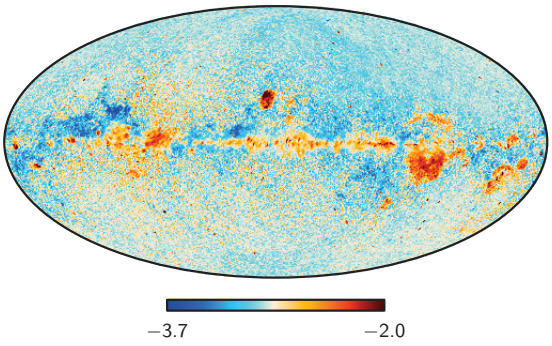


(b) CO amplitude at 100 GHz

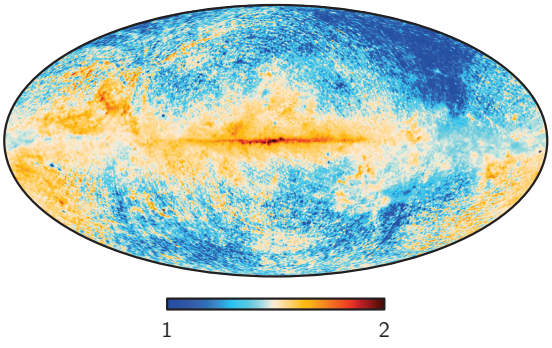


(c) Thermal dust amplitude at 353 GHz

Fig. 15: Posterior mean foreground amplitude maps derived from the low-resolution analysis. From top to bottom are shown the low-frequency, CO and thermal dust emission maps.



(a) Low-frequency index



(b) Dust emissivity

Fig. 16: Posterior mean spectral parameter maps derived from the low-resolution analysis. The top panel shows the power law index of the low-frequency component, and the bottom panel shows the emissivity index of the one-component thermal dust model. Note that the systematic error due to monopole and dipole uncertainties is significant for the dust emissivity in regions with a low thermal dust amplitude.

low angular resolution, followed by a generalized least-squares solution for component amplitudes at high resolution.

8.1. Data selection and processing

We only use the seven lowest *Planck* frequencies, from 30 to 353 GHz. The two highest channels have significantly different systematic properties than the lower frequency bands, for instance concerning calibration, ZLE, and noise correlations, and they are more relevant to thermal dust and CIB studies than to the present CMB analysis. Studies of specific foregrounds (CO, thermal dust, CIB etc.) using all *Planck* frequencies as well as ancillary data are discussed in companion papers Planck Collaboration VI (2013), Planck Collaboration XXI (2013), Planck Collaboration (2013b), and additional future publications will consider extensions to AME, synchrotron and free-free emission.

In order to obtain unbiased estimates of the spectral parameters across all frequency bands, each map is downgraded from its native resolution to a common angular resolution of 40′ and repixelized at $N_{\rm side} = 256$, a limit imposed by the LFI 30 GHz channel. Once the spectral indices have been determined, we re-estimate the component amplitudes at native *Planck* resolution (see Appendix A).

Although the smoothing operation introduces noise correlations between pixels, we model the noise of the smoothed maps as uncorrelated white noise with an effective standard deviation, $\sigma(p)$, for each pixel p. This approximation does not bias the final solution, because the analysis is performed independently for each pixel. However, it is important to note that correlations between pixels are not taken into account in this analysis. The ef-

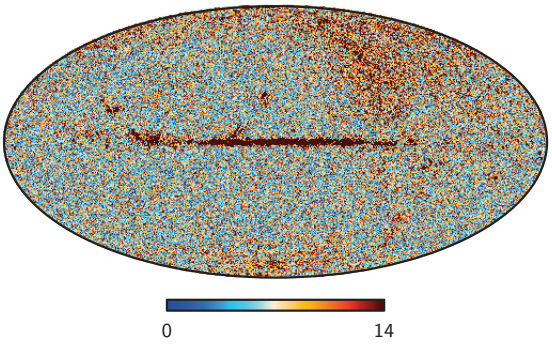


Fig. 17: χ^2 per pixel for the joint CMB and foreground analysis. The expected value for an acceptable fit is 7, corresponding to the number of frequency bands used in this analysis. The pixels with high values can be classified into two types, due to either modelling errors (i.e., high residuals in the Galactic plane) or to un-modelled correlated noise (i.e., stripes crossing through low dust emission regions).

fective noise uncertainty, $\sigma(p)$, is estimated using realistic noise simulations downgraded in the same way as the data. The measured instrumental bandpasses are taken into account by integrating the emission laws over the bandpass for each component at each Monte Carlo step in the analysis.

The monopole (zero-point) of each frequency map is not constrained by *Planck*, but is rather determined by post-processing, and associated with a non-negligible uncertainty (see Table 5 of Planck Collaboration I 2013). In addition, each frequency map includes a significant monopole contribution from isotropic extragalactic sources and CIB fluctuations not traced by local Galactic structure, ranging from less than about 10–20 μK at 70 GHz to several hundreds of μK at 353 GHz. Finally, the effective dipole in each map is associated with significant uncertainty due to the large kinematic CMB dipole. In order to prevent these effects from introducing modelling errors during component separation, they must be fit either prior to or jointly with the Galactic parameters. Unfortunately, when allowing free spectral parameters per pixel, there is a near-perfect degeneracy among the offsets, the foreground amplitudes and the spectral indices, and in order to break this degeneracy, it is necessary to reduce the number of spectral degrees of freedom.

We adopt the method described by Wehus et al. (2013) for this purpose, which has the additional advantage of making minimal assumptions about the foreground spectra. In short, this method uses linear regression between data from CMB-subtracted maps evaluated on pixels falling within each large $N_{\rm side} = 8$ pixel to estimate the relative offsets, m_1 and m_2, between any two maps at each position on the sky. Each regression provides a constraint of the form $m_1 = a m_2 + b$, where a and b are the slope and offset, respectively, and where each value of m_i consists of the sum of both a monopole and a dipole term evaluated at that position. The individual monopoles and dipoles can then be reconstructed by measuring a and b in different regions of the sky, exploiting spatial variations in spectral indices, and solving jointly for two monopoles and dipoles, including constraints from all positions. To minimize degeneracies, a positivity prior is imposed on the fit, such that statistically significant negative pixels are heavily penalized. For 44 and 70 GHz, we retain the dipole values determined during the mapmaking process, and do not attempt to fit them.

The resulting complete set of monopole and dipole values is listed in Table 2. As a cross-check, we performed a dedicated `Commander` run in which we fitted for the dipole at 353 GHz, together with the foreground amplitudes and spectral indices, and only found sub-μK differences. This channel is by far the most problematic in our data set in terms of offset determination, because of the very bright dust emission at this frequency. As a result, there is a large relative uncertainty between the zero-level of the dust amplitude map and the 353 GHz channel offset not accounted for in the following analyses. However, the sum of the two terms is well determined, and a potential error in either therefore does not compromise the quality of the other signal components (e.g., CMB and low-frequency components). A similar comment applies between the offset at 30 GHz and the zero-level of the low-frequency component, although at a significantly lower level.

8.2. Component models and priors

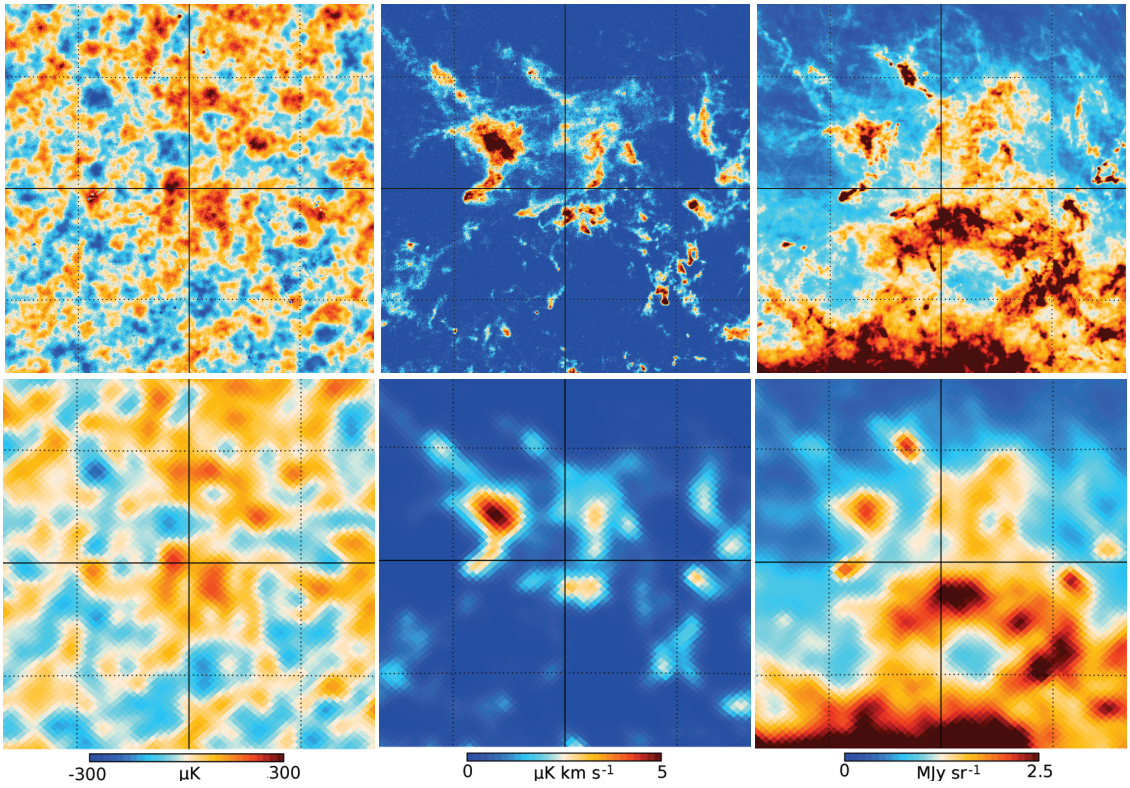
Our model for the low-resolution CMB analysis includes four independent physical components: CMB; "low-frequency" emission; CO emission; and thermal dust emission. It can be written schematically in the form

$$s_\nu(p) = A_{\rm CMB}(p) + A_{\rm lf}(p)\left(\frac{\nu}{\nu_{0,\rm lf}}\right)^{\beta_{\rm lf}(p)} + \\ + A_{\rm CO}(p) f_{\nu,\rm CO} + A_{\rm d}(p)\frac{e^{\frac{h\nu_{0,\rm d}}{kT_{\rm d}(p)}} - 1}{e^{\frac{h\nu}{kT_{\rm d}(p)}} - 1}\left(\frac{\nu}{\nu_{0,\rm d}}\right)^{\beta_{\rm d}(p)+1}, \quad (1)$$

where $A_i(p)$ denotes the signal amplitude for component i at pixel p, $\nu_{0,i}$ is the reference frequency for each component, and ν refers to frequency. (Note that for readability, integration over bandpass, as well as unit conversions between antenna, flux density and thermodynamic units, is suppressed in this expression.) Thus, each component is modelled with a simple frequency spectrum parameterized in terms of an amplitude and a small set of free spectral parameters (a power-law index for the low-frequency component, and an emissivity index and temperature for the thermal dust component); no spatial priors are imposed. One goal of the present analysis is to understand how well this simple model captures the sky signal in terms of effective components over the considered frequency range, and we exploit the FFP6 simulation (see Sect. 4) for this purpose.

In order to take into account the effect of bandpass integration, each term in the above model is evaluated as an integral over the bandpass as described in Sect. 3 of Planck Collaboration IX (2013), and converted internally to thermodynamic units. Accordingly, the reference frequencies in Eq. 1 are computed as effective integrals over the bandpass, such that the amplitude map, $A_i(p)$, corresponds to the foreground map observed by the reference detector, i.e., after taking into account the bandpass. In order to minimize degeneracies between the different signal components, the reference band for a given component is set to the frequency at which its relative signal-to-noise ratio is maximized.

The foreground model defined in Eq. 1 is motivated by prior knowledge about the foreground composition over the CMB frequencies as outlined in Sect. 2, as is our choice of priors. In



(a) CMB amplitude (b) CO amplitude at 100 GHz (c) Thermal dust amplitude at 353 GHz

Fig. 18: Comparison of the high-resolution `Ruler` (top) and low-resolution `Commander` (bottom) amplitude maps for a particularly strong CO complex near the Fan region; the maps are centred on Galactic coordinates $(l, b) = (110°, 15°)$, and the grid spacing is 5°. Columns show, left to right: the CMB amplitude; the CO amplitude at 100 GHz and the thermal dust amplitude at 353 GHz.

Table 2: Estimated monopoles and dipoles in Galactic coordinates, all measured in thermodynamic μK. Errors are estimated by bootstrapping, and do not account for correlated errors across frequencies. In particular, the 353 GHz monopole uncertainty is dominated by systematic errors not included in these estimates. Note that the dipoles at 44 and 70 GHz are fixed at the values determined in the mapmaking.

Frequency [GHz]	Monopole [μK]	X dipole [μK]	Y dipole [μK]	Z dipole [μK]
30	8 ± 2	-4 ± 3	-6 ± 2	6 ± 1
44	2 ± 1	0 ± 0	0 ± 0	0 ± 0
70	15 ± 1	0 ± 0	0 ± 0	0 ± 0
100	15 ± 1	2 ± 1	5 ± 1	-5 ± 1
143	33 ± 1	2 ± 1	7 ± 1	-6 ± 1
217	86 ± 1	2 ± 1	11 ± 2	-10 ± 2
353	414 ± 4	11 ± 10	52 ± 12	-37 ± 8

addition to the Jeffreys prior[5] (Eriksen et al. 2008), we adopt Gaussian priors on all spectral parameters with centre values and widths attempting to strike a balance between prior knowledge and allowing the data to find the optimal solution. Where needed, we have also run dedicated analyses, either including particular high signal-to-noise ratio subsets of the data or using a lower resolution parameterization to increase the effective signal-to-noise in order to inform our prior choices. We now consider each foreground component in turn, and note in passing that the CMB component, by virtue of being a blackbody signal, is given by a constant in thermodynamic temperature units.

We approximate the low-frequency component by a straight power law in antenna temperature with a free spectral index per pixel, and adopt a prior of $\beta = -3 \pm 0.3$ (this is the index in terms of brightness temperature). This choice is determined by

[5] The purpose of the Jeffreys prior is to normalize the parameter volume relative to the likelihood, such that the likelihood becomes so-called "data-translated", i.e., invariant under re-parameterizations.

noting that the prior is in practice only relevant at high Galactic latitudes where the signal-to-noise ratio is low and the dominant foreground component is expected to be synchrotron emission; in the signal-dominated and low-latitude AME and free-free regions, the data are sufficiently strong to render the prior irrelevant. For validation purposes, we have also considered minor variations around this prior, such as $\beta = -2.9 \pm 0.3$ and $\beta = -3.05 \pm 0.2$, finding only small differences in the final solutions. The reference band for the low-frequency component is set to 30 GHz, where the low-frequency foreground signal peaks. The final low-frequency amplitude map is provided in units of thermodynamic microkelvin.

The CO emission is modelled in terms of a single line ratio for each frequency. Specifically, the CO amplitude is normalized to the 100 GHz band, and defined in units of $\mu\mathrm{K\,km\,s^{-1}}$ (Planck Collaboration XIII 2013). The amplitude at other frequencies is determined by a single multiplicative factor relative to this, with a numerical value of 0.595 at 217 GHz and 0.297 at 353 GHz; all other frequencies are set to zero. These values are obtained from a dedicated CO analysis that includes only high signal-to-noise ratio CO regions covering a total of 0.5% of the sky. The derived values are in good agreement with those presented by Planck Collaboration XIII (2013).

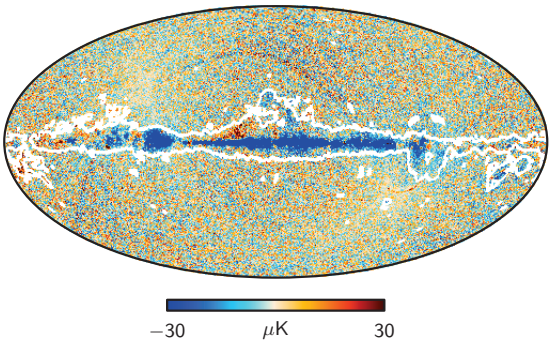
Thermal dust emission is modelled by a one-component modified blackbody emission law with a free emissivity spectral index, β_d, and dust temperature, T_d, per pixel. However, since we only include frequencies below 353 GHz, the dust temperature is largely unconstrained in our fits, and we therefore impose a tight prior around the commonly accepted mean value of $T_\mathrm{d} = 18 \pm 0.05$ K. The only reason we do not fix it completely to 18 K is to allow for modelling errors near the Galactic centre. The dust emissivity prior is set to $\beta_\mathrm{d} = 1.5 \pm 0.3$, where the mean is determined by a dedicated run fitting for a single best-fit value for the high-latitude sky, where the prior is relevant. The reference band for the thermal dust component is 353 GHz, and the final map is provided in units of megajansky per steradian.

8.3. Results and validation

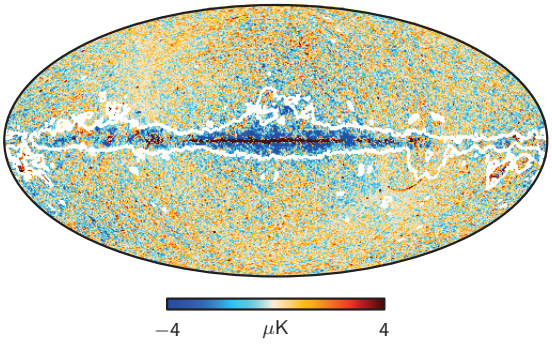
The output of the Bayesian component separation algorithm is a set of samples drawn from the joint posterior distribution of the model parameters, as opposed to a single well-defined value for each. For convenience, we summarize this distribution in terms of posterior mean and standard deviation maps, computed over the sample set, after rejecting a short burn-in phase. The goodness-of-fit is monitored in terms of the χ^2 per pixel. Although convenient, it is, however, important to note that this description does not provide a comprehensive statistical representation of the full posterior distribution, which is intrinsically non-Gaussian. One should be careful about making inferences in the low signal-to-noise regime based on this simplified description.

The low-resolution `Commander` posterior mean amplitude maps are shown in Fig. 15 for the low-frequency, CO, and thermal dust components, and the spectral index maps in Fig. 16. The associated χ^2 map is plotted in Fig. 17. Note that because we are sampling from the posterior instead of searching for the maximum-likelihood point, the expected number of degrees of freedom is equal to $N_\mathrm{band} = 7$ in this plot, not $N_\mathrm{band} - N_\mathrm{par}$.

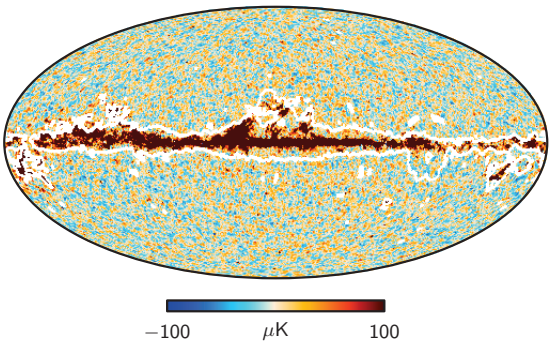
Figure 18 compares the high-resolution `Ruler` solution to the low-resolution `Commander` solution for CMB, CO and thermal dust on a particularly strong CO complex near the Fan region, centred on Galactic coordinates $(l, b) = (110°, 15°)$.



(a) Low-frequency component residual at 30 GHz



(b) CO residual at 100 GHz



(c) Thermal dust residual at 353 GHz

Fig. 19: Amplitude residual maps, $A_\mathrm{out} - A_\mathrm{in}$, computed blindly from the FFP6 simulation. The panels show (from top to bottom) the low-frequency residual at 30 GHz, the CO residual at 100 GHz and the thermal dust residual at 353 GHz. All units are thermodynamic μK. The white lines indicate the boundary of the `Commander` likelihood analysis mask, removing 13% of the sky.

Several features can be seen here, foremost of which is that the Galactic plane is strikingly obvious, with χ^2 values exceeding 10^4 for seven degrees-of-freedom in a few pixels. This is not surprising, given the very simplified model at low frequencies (i.e., a single power law accounting for AME, synchrotron, and free-free emission), as well as the assumption of a nearly

constant dust temperature of 18 K. Second, there is an extended region of moderately high χ^2 roughly aligned with a great circle going through the Ecliptic South Pole, indicating the presence of correlated noise in the scanning rings not accounted for in our white noise model.

Based on these, and other considerations, it is clear that parts of the sky must be masked before proceeding to CMB power spectrum and likelihood analyses. This masking process is discussed at greater length in Planck Collaboration XV (2013), and results in different masks for specific applications. The goal of our present discussion is to evaluate the adequacy of the mask adopted for low-ℓ likelihood analysis (L87), which is based on the fits presented here. This mask removes 13% of the sky, and is derived from a combination of χ^2 and component amplitude thresholding.

For validation purposes, we analyse the simulations described in Section 4 in the same way as the real data, including monopole and dipole determination, CO line ratio estimation and spectral index estimation. Individual component maps at each observed frequency are available from the simulation process, and used for direct comparison with the reconstructed products.

In Fig. 19 we show the differences between the recovered and input component maps at their respective reference frequencies. The boundary of the 13% `Commander` mask is traced by the white contours, and a best-fit monopole and dipole have been subtracted from each difference map. All difference maps are shown in units of thermodynamic μK. The top panel of Fig. 20 gives the error histograms outside the masked region for each component, normalized to the respective estimated standard deviation; if the recovered solution has both correct mean and standard deviation, these histograms should match a Gaussian distribution with zero mean and unit variance, indicated by the dashed black line. Conversely, a significant bias would be visible as a horizontal shift in this plot, while under-estimation of the errors would result in too wide a distribution and vice versa. The bottom panel shows the fractional error (i.e., the error divided by the true input value) for all pixels with signal above $5\,\sigma$; the fractional error is not a useful quantity for noisier signals.

The difference maps in Fig. 19 display significant errors in the Galactic plane. For the low-frequency component, the residuals are dominated by free-free emission, while for thermal dust the dominant contaminant is CO emission. However, outside the mask the residuals are small, and, at least for the low-frequency and CO components, the spatial characteristics appear similar to instrumental noise. This is more clear in the histograms shown in the top panel of Fig. 20; the mean and standard deviations are $\delta_{\rm lf} = 0.01 \pm 1.12$, $\delta_{\rm CO} = 0.00 \pm 0.87$, and $\delta_{\rm td} = 0.00 \pm 2.01$, respectively, for the low-frequency, CO and thermal dust components. There is no evidence of bias outside the mask in any component, and the error estimates are accurate to 12 and 13 % for the low-frequency and CO components. Note, though, that the estimated error for the CO component is actually larger than the true uncertainty, suggesting that the white noise approximation for the 100 GHz channel overestimates the true noise. This can occur if the correlated instrumental noise is important in regions where there is no significant CO emission. Locally re-scaling the white noise to account for spatially varying correlated noise would correct this effect.

For the thermal dust component, on the other hand, the error is underestimated by a factor of 2. The explanation for this is most easily seen from the lower panel of Fig. 19. This map is dominated by isotropic CIB fluctuations, rather than instrumental noise. Because these fluctuations have a slightly different spectrum than the dominant Galactic dust emission, and the model does not account for a separate CIB component, the error on the Galactic component is underestimated. When using the Galactic map presented here for detailed analysis near the noise limit, taking into account these residual fluctuations is essential, and the effective noise per pixel should be increased by a factor of 2.

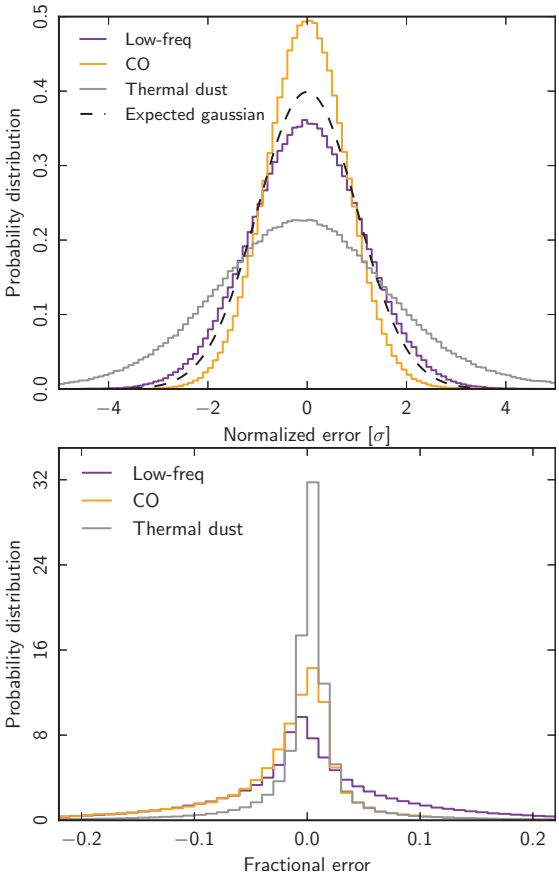


Fig. 20: Error validation for component amplitudes, evaluated from the FFP6 simulation. The upper panel shows histograms of the normalized errors $\delta = (A_{\rm out} - A_{\rm in})/\sigma_{\rm out}$ for the three foreground components and including all pixels outside the `Commander` likelihood analysis mask. The lower panel shows histograms of the fractional error $f \equiv (A_{\rm out} - A_{\rm in})/A_{\rm in}$ for pixels with a foreground detection level above $5\,\sigma$. No evidence of significant bias is observed for any component, and the uncertainty estimates for the low-frequency and CO components are accurate to about 12 %; the thermal dust uncertainty is underestimated by a factor of 2 due to the presence of unmodelled fluctuations.

As clearly seen in Fig. 19, the residuals inside the mask are highly significant in a strict statistical sense. However, as seen in the bottom panel of Fig. 20, they are relatively small in terms of fractional errors. Specifically, the three histograms have means

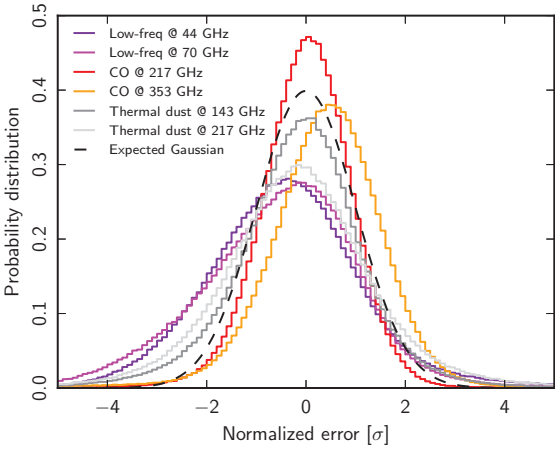


Fig. 21: Validation of spectral parameters for low-frequency foregrounds, thermal dust, and CO emission, evaluated from the FFP6 simulation. Each histogram shows the error distribution at the two leading sub-dominant frequencies in the form of the normalized errors $\delta = (A_{\rm out}(\nu) - A_{\rm in}(\nu))/\sigma_{\rm out}(\nu)$ for all pixels outside the `Commander` likelihood analysis mask, where $A_{\rm out}(\nu)$ is the predicted foreground amplitude at frequency ν given the estimated amplitude and spectral parameters, and $\sigma_{\rm out}(\nu)$ is the corresponding standard deviation computed over the sample set.

and standard deviations of $f_{\rm lf} = 0.00 \pm 0.10$, $f_{\rm CO} = -0.03 \pm 0.10$, and $f_{\rm fd} = 0.00 \pm 0.06$, respectively, for the low-frequency, CO and thermal dust components. The largest bias is observed for the CO component, for which the absolute amplitude is biased by 3 %. The bias in the low-frequency and thermal dust components is negligible, and the fractional uncertainties are 10 and 6 %, respectively. This confirms that approximating the sum of the three low-frequency components by a single power-law over the *Planck* frequency bands is reasonable; if modelling errors dominated, one would expect to see a significant bias in the resulting amplitude.

In order to validate the spectral parameters, we show in Fig. 21 histograms of the normalized residuals for each foreground component evaluated at its two leading sub-dominant frequencies (i.e., at 44 and 70 GHz for the low-frequency component; at 217 and 353 GHz for the CO component; and at 143 and 217 GHz for the thermal dust component). The means and standard deviations of these distributions are: $\delta_{\rm lf}(44\,{\rm GHz}) = -0.41 \pm 1.98$ and $\delta_{\rm lf}(70\,{\rm GHz}) = -0.34 \pm 2.04$ for the low-frequency component; $\delta_{\rm CO}(217\,{\rm GHz}) = 0.10 \pm 0.84$ and $\delta_{\rm CO}(353\,{\rm GHz}) = 0.51 \pm 1.00$ for the CO component; and $\delta_{\rm td}(143\,{\rm GHz}) = -0.02 \pm 1.53$ and $\delta_{\rm td}(217\,{\rm GHz}) = -0.13 \pm 1.87$ for the thermal dust component. As expected, the effect of modelling errors is more significant at the sub-dominant frequencies than at the pivot frequencies, when measured in terms of statistical uncertainties, since the foreground signal is weaker and the confusion with the other components relatively larger. Nevertheless, we see that the absolute bias is at most $0.5\,\sigma$ for the CO component at 353 GHz, while the thermal dust bias is negligible even at 143 GHz. The estimated uncertainties are generally underestimated by up to a factor of two due to these modelling errors.

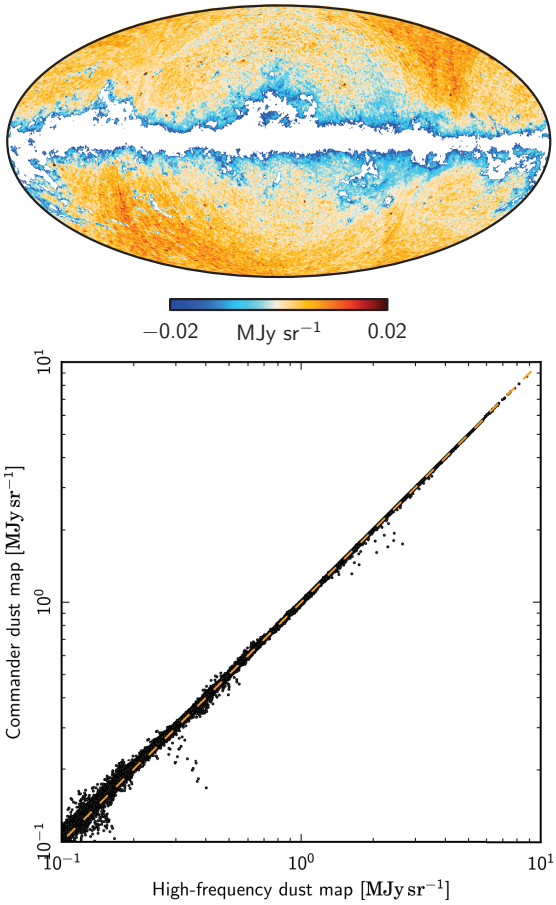


Fig. 22: Top: Difference map between the estimated thermal dust amplitude at 353 GHz derived by Planck Collaboration (2013b) and the low-resolution dust map presented here, both smoothed to 40′. A monopole and dipole term outside the `Commander` mask has been removed. The former includes only high-frequency observations (the *Planck* 353, 545 and 857 GHz channels and observations at 100 μm), while the one derived here only uses low-frequency data (*Planck* 30–353 GHz). Note that the colour scale ranges between −0.02 and 0.02 MJy sr^{-1} in this plot, whereas it ranges from 0 to 2.5 MJy sr^{-1} in the bottom panel of Fig. 15. Masked regions indicate pixels with a CO amplitude at 100 GHz larger than 1 μK km s^{-1}. Bottom: T–T plot between the same two maps.

Finally, the efficiency of the adopted foreground model for CMB analysis is quantified in Appendix E in terms of power spectrum residuals and cosmological parameter estimation.

To summarize, we find that the simplified model, defined by Eq. 1, provides a good fit to the realistic FFP6 simulation for most of the sky. Absolute residuals are small, and the amplitude uncertainty estimates are accurate to around 12 %, except for the thermal dust component for which unmodelled CIB fluctuations

are important. Further, we find that the real *Planck* data behave both qualitatively and quantitatively very similarly to the FFP6 simulation, suggesting that this approach also performs well on the real sky.

8.4. Interpretation and comparison with other results

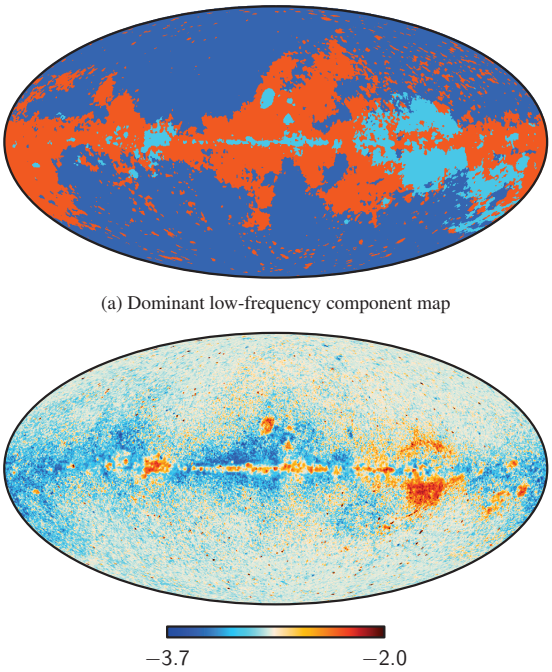
The maps shown in Figs. 15 and 16 provide a succinct summary of the average foreground properties over the *Planck* frequency range. We now consider their physical interpretation and compare them to products from alternative methods.

First, the top panel of Fig. 22 shows a difference map between the dust map at 353 GHz derived in the present paper and one determined from only the three highest *Planck* frequencies and the 100 μm IRIS map by Planck Collaboration (2013b), shown in flux density units, after removing a small monopole and dipole difference. This map is to be compared to the corresponding dust amplitude map in the bottom panel of Figure 15; note that the colour range varies from -0.02 to 0.02 MJy sr^{-1} in the difference plot, and between 0 and 2.5 MJy sr^{-1} in the amplitude plot. Thus, despite the very different data sets and methods, we see that the two reconstructions agree to very high accuracy outside the Galactic plane. Inside the Galactic plane, the differences are dominated by residuals due to different CO modelling, seen as solid blue colours in Fig. 22; however, even in this region the differences are smaller than 5 % of the amplitude. The bottom panel shows a corresponding T–T plot between the two maps, excluding any pixel for which the Commander CO amplitude at 100 GHz is larger than 1 μK km s^{-1}. The two maps agree to 0.2% in terms of best-fit amplitude.

From Fig. 16, we see that the dust emissivity ranges between 1.3 and 1.7 for most of the sky; considering only the pixels with a posterior distribution width that is a third of the prior width (i.e., $\sigma(\beta_{\rm d}) < 0.1$), we find a mean value of 1.49. The two exceptions are a large region of shallow indices northeast of the Galactic centre, and steep indices near the Galactic plane. The former region corresponds to a part of the sky with low dust emission, where we expect the spectral index to be sensitive to both monopole and dipole residuals, as well as instrumental systematics, such as correlated $1/f$ noise. The latter appears to be particularly pertinent here because the shallow index region at least partially traces the *Planck* scanning strategy; as a result, the systematic error on the spectral index in this region is considerable. The main systematic uncertainty connected to the region of steep indices around the Galactic plane is confusion with CO emission.

The CO map shown in Figs. 15 and 18 is discussed in greater detail in Planck Collaboration XIII (2013). A distinct advantage of this particular solution over available alternatives is its high signal-to-noise ratio per pixel, which is achieved by reducing all information into a single value per pixel. Consequently, this map serves as a unique tool for follow-up CO observations. However, the assumption of a constant line ratio over the full sky may lead to a significant systematic uncertainty on CO amplitude per pixel. This possibility is investigated in a forthcoming work (Planck Collaboration 2013a).

Finally, the spectral index map for the low-frequency component shown in Fig. 16 can be used to determine the dominant low-frequency component (synchrotron, free-free or AME) as a function of position on the sky. To illustrate this connection, we once again take advantage of the FFP6 simulation for which we know the amplitude of each low-frequency component per pixel. In the top panel of Fig. 23, we use this information to make a "dominant component map"; dark blue indicates that synchrotron emission is strongest at a given pixel, light blue that free-free is strongest, and orange that spinning dust (AME) dominates. In the bottom panel, we show our derived power-law index map from the same simulation. As expected, the correspondence between the power-law index and the dominant component is very strong, implying that the spectral index map can be used to trace the individual components. In particular, we see that an index below about -3.3 reflects the presence of a component consistent with a spinning dust model peaking below 20 GHz over the *Planck* frequency range, while an index higher than around -2.3 signals the importance of free-free emission. Intermediate values typically indicate synchrotron emission, although it should be noted that the signal-to-noise ratio at very high latitudes is low and the results are therefore prior-driven in these regions.



(a) Dominant low-frequency component map

(b) Low-frequency component power-law index

Fig. 23: Top: Dominant foreground component per pixel at 30 GHz in the FFP6 simulation. Dark blue indicates that synchrotron emission is the strongest component at 30 GHz, light blue indicates that free-free dominates, and orange indicates that spinning dust (AME) is the strongest component. Bottom: The recovered low-frequency power-law index derived from the same simulation.

Returning to the spectral index map shown in Fig. 16, we see a good correspondence between the real data and the simulation. Features present in the simulation also appear in the data. For instance, we see that the spectral index in the so-called Fan Region (i.e., near Galactic coordinates $(l, b) = (90^\circ, 20^\circ)$) is low in both cases, and this alone provides strong evidence for the presence of AME. Further, the AME spectral index is consistent with the spinning dust interpretation. This power-law index map may be used to identify particular AME regions for follow-up observa-

tions. Finally, we note that regions known for strong free-free emission, such as the Gum Nebula or Zeta Ophiuchi, have spectral indices close to −2.1 or −2.2, as expected.

9. Conclusions

We have tested four component separation algorithms on the *Planck* frequency maps to produce clean maps of the CMB anisotropies over a large area of sky. These CMB maps are used for studies of statistics and isotropy (Planck Collaboration XXIII 2013), primordial non-Gaussianity (Planck Collaboration XXIV 2013), gravitational lensing (Planck Collaboration XVII 2013), the ISW effect (Planck Collaboration XIX 2013), cosmic geometry and topology (Planck Collaboration XXVI 2013), searches for cosmic defects from primordial phase transitions (Planck Collaboration XXV 2013), as well as an integral part of the low-ℓ *Planck* likelihood (Planck Collaboration XV 2013). Two of the methods, one using internal foreground templates (`SEVEM`) and the other an ILC in needlet space (`NILC`), are non-parametric, extracting the CMB map by minimizing the variance of the total contamination. The other two methods fit models of the foregrounds to clean the CMB of their emission. One fits a parametric model in real space (`C-R`) and one fits a non-parametric in the harmonic domain (`SMICA`).

All four methods have been demonstrated to work well both on real and simulated data, and to yield consistent results. Nevertheless, there are differences between the methods, making them more or less suitable for specific applications. For instance, `Commander-Ruler` allows a joint parametric foreground estimation and CMB power spectrum estimation, with full propagation of foreground uncertainties to cosmological parameters, but is limited to a lower angular resolution than the other codes. This method has therefore been selected for the low-ℓ *Planck* likelihood (Planck Collaboration XV 2013) and to produce astrophysical component maps (Sect. 8), while it is sub-optimal for applications requiring full angular resolution, e.g., gravitational lensing reconstruction or estimation of primordial non-Gaussianity. For these purposes, we use the three higher-resolution maps. We take `SMICA` to be the leading method, based on its superior performance on the FFP6 simulation, where it has be shown to have the lowest residual foreground contamination at large scales and to preserve primordial non-Gaussianity. When subjecting foreground-cleaned *Planck* maps to scientific analysis, we use the other two or three maps, as appropriate, to assess the uncertainties inherent in the choice of methods and the assumptions they make. Indeed, this is the main purpose for presenting four different CMB solutions to the general community.

The CMB anisotropies are robustly recovered over a large fraction (73 %) of the sky and down to small angular scales, reaching to multipoles $\ell \approx 2000$. We characterize the CMB maps with angular power spectra and cosmological parameter constraints. Parameter constraints from these maps are consistent with those from the *Planck* likelihood function based on cross-spectra and large sky cuts (Planck Collaboration XV 2013). This agreement supports the robustness of both our component separation methodology and cosmological parameter constraints.

The real-space parametric fits of `Commander-Ruler` enable us to characterize the diffuse Galactic foregrounds. We parameterize them with a low-frequency power-law component, representing the sum of synchrotron, free-free, and AME emission, a high-frequency modified blackbody spectrum describing thermal dust emission, and a molecular CO component. Using only the *Planck* data from 30 to 353 GHz, we fit for the amplitude and spectral parameters of the three foregrounds and the CMB simultaneously at each pixel of a 40-arcmin resolution map. The spectral parameters are the low-frequency component power-law exponent and the modified blackbody emissivity power-law exponent; the CO line ratios are spatially fixed. These parameters give us the source mixing matrix, which we then use in a direct inversion to deduce the component amplitudes at higher resolution. Through Gibbs sampling, we obtain realizations drawn from the full posterior distribution of possible foreground and CMB solutions, giving us a powerful ability to statistically characterize our results.

Our in-depth analysis of the recovered CMB anisotropies is unprecedented for component separation studies, concerning both the accuracy of cosmological parameter constraints, and studies of early Universe physics and structure formation through gravitational lensing. On the other hand, the complex nature of the foreground emission over such a large frequency range limits us to the use of relatively simple methods when analysing *Planck* data alone. An extensive study in combination with other probes of Galactic foregrounds will be presented in forthcoming papers. In particular, the separation of individual components at low frequencies requires the use of ancillary data, for example, from the Wilkinson Microwave Anisotropy Probe (*WMAP*) and radio surveys.

Acknowledgements. The development of Planck has been supported by: ESA; CNES and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MICINN, JA and RES (Spain); Tekes, AoF and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); PRACE (EU). A description of the Planck Collaboration and a list of its members, including the technical or scientific activities in which they have been involved, can be found at `http://www.sciops.esa.int/index.php?project=planck&page=Planck_Collaboration`. The authors acknowledge the support provided by the Advanced Computing and e-Science team at IFCA. This work made use of the COSMOS supercomputer, part of the STFC DiRAC HPC Facility. Some of the results in this paper have been derived using the `HEALPix` package.

References

Ali-Haïmoud, Y., Hirata, C. M., & Dickinson, C. 2009, MNRAS, 395, 1055
Banday, A. J., Dickinson, C., Davies, R. D., Davis, R. J., & Górski, K. M. 2003, MNRAS, 345, 897
Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2003, ApJS, 148, 97
Benoit-Lévy, A., Déchelette, T., Benabed, K., et al. 2013, A&A, accepted
Betoule, M., Pierpaoli, E., Delabrouille, J., Le Jeune, M., & Cardoso, J.-F. 2009, A&A, 503, 691
Bonaldi, A., Ricciardi, S., Leach, S., et al. 2007, MNRAS, 382, 1791
Broadbent, A., Osborne, J. L., & Haslam, C. G. T. 1989, MNRAS, 237, 381
Cardoso, J.-F., Martin, M., Delabrouille, J., Betoule, M., & Patanchon, G. 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 735
Casaponsa, B., Barreiro, R. B., Curto, A., Martínez-González, E., & Vielva, P. 2011, MNRAS, 411, 2019
Davies, R. D., Dickinson, C., Banday, A. J., et al. 2006, MNRAS, 370, 1125
Davies, R. D., Watson, R. A., & Gutierrez, C. M. 1996, MNRAS, 278, 925
de Oliveira-Costa, A., Tegmark, M., Davies, R. D., et al. 2004, ApJ, 606, L89
Delabrouille, J., Betoule, M., Melin, J.-B., et al. 2012, ArXiv e-prints
Delabrouille, J., Cardoso, J.-F., Le Jeune, M., et al. 2009, A&A, 493, 835
Dickinson, C., Davies, R. D., & Davis, R. J. 2003, MNRAS, 341, 369
Dobler, G. & Finkbeiner, D. P. 2008, ApJ, 680, 1222
Draine, B. T. & Lazarian, A. 1998, ApJ, 508, 157
Eriksen, H. K., Dickinson, C., Lawrence, C. R., et al. 2006, ApJ, 641, 665
Eriksen, H. K., Huey, G., Saha, R., et al. 2007, ApJ, 656, 641
Eriksen, H. K., Jewell, J. B., Dickinson, C., et al. 2008, ApJ, 676, 10
Eriksen, H. K., O'Dwyer, I. J., Jewell, J. B., et al. 2004, ApJS, 155, 227
Fendt, W. & Wandelt, B. 2008, in APS April Meeting Abstracts, J8004
Fernández-Cobos, R., Vielva, P., Barreiro, R. B., & Martínez-González, E. 2012, MNRAS, 420, 2162
Finkbeiner, D. P. 2004, ApJ, 614, 186
Finkbeiner, D. P., Davis, M., & Schlegel, D. J. 1999, ApJ, 524, 867
Finkbeiner, D. P., Langston, G. I., & Minter, A. H. 2004, ApJ, 617, 350

Ghosh, T., Banday, A. J., Jaffe, T., et al. 2012, MNRAS, 422, 3617
Gold, B., Odegard, N., Weiland, J. L., et al. 2011, ApJS, 192, 15
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Haslam, C., Stoffel, H., Salter, C. J., & Wilson, W. E. 1982, Astronomy and Astrophysics Supplement Series, 47, 1
Hivon, E., Górski, K. M., Netterfield, C. B., P., C. B., & et al. 2002, Astrophys.J., 567
Hoang, T. & Lazarian, A. 2012, Advances in Astronomy, 2012
Jewell, J., Levin, S., & Anderson, C. H. 2004, ApJ, 609, 1
Kogut, A. 1996, in Bulletin of the American Astronomical Society, Vol. 28, American Astronomical Society Meeting Abstracts, 1295
Komatsu, E., Spergel, D. N., & Wandelt, B. D. 2005, ApJ, 634, 14
Lagache, G. 2003, A&A, 405, 813
Leach, S. M., Cardoso, J., Baccigalupi, C., et al. 2008, A&A, 491, 597
Leitch, E. M., Readhead, A. C. S., Pearson, T. J., & Myers, S. T. 1997, ApJ, 486, L23
Martınez-Gonzalez, E., Diego, J. M., Vielva, P., & Silk, J. 2003, MNRAS, 345
Miville-Deschênes, M.-A., Ysard, N., Lavabre, A., et al. 2008, A&A, 490, 1093
Okamoto, T. & Hu, W. 2003, Phys. Rev., D67, 083002
Patanchon, G., Cardoso, J.-F., Delabrouille, J., & Vielva, P. 2005, MNRAS, 364, 1185
Pietrobon, D., Górski, K. M., Bartlett, J., et al. 2012, ApJ, 755, 69
Planck Collaboration. 2012, Planck intermediate results: Dust emission at millimetre wavelenghts in the Galactic Plane (in preparation)
Planck Collaboration. 2013a, In preparation
Planck Collaboration. 2013b, Submitted to A&A
Planck Collaboration ES. 2013, The Explanatory Supplement to the Planck 2013 results (ESA)
Planck Collaboration I. 2013, Submitted to A&A
Planck Collaboration II. 2013, Submitted to A&A
Planck Collaboration Int. IX. 2013, Submitted to A&A
Planck Collaboration Int. XII. 2013, Submitted to A&A
Planck Collaboration IX. 2013, Submitted to A&A
Planck Collaboration VI. 2013, Submitted to A&A
Planck Collaboration VIII. 2013, Submitted to A&A
Planck Collaboration X. 2013, Submitted to A&A
Planck Collaboration XIII. 2011, A&A, 536, A13
Planck Collaboration XIII. 2013, Submitted to A&A
Planck Collaboration XIX. 2013, Submitted to A&A
Planck Collaboration XV. 2013, Submitted to A&A
Planck Collaboration XVI. 2013, Submitted to A&A
Planck Collaboration XVII. 2013, Submitted to A&A
Planck Collaboration XVIII. 2011, A&A, 536, A18
Planck Collaboration XVIII. 2013, Submitted to A&A
Planck Collaboration XX. 2011, A&A, 536, A20
Planck Collaboration XXI. 2013, Submitted to A&A
Planck Collaboration XXII. 2013, Submitted to A&A
Planck Collaboration XXIII. 2013, Submitted to A&A
Planck Collaboration XXIV. 2013, Submitted to A&A
Planck Collaboration XXIX. 2013, Submitted to A&A
Planck Collaboration XXV. 2013, Submitted to A&A
Planck Collaboration XXVI. 2013, Submitted to A&A
Planck Collaboration XXVIII. 2013, Submitted to A&A
Platania, P., Burigana, C., Maino, D., et al. 2003, A&A, 410, 847
Reich, P. & Reich, W. 1988, A&AS, 74, 7
Rocha, G., Contaldi, C. R., Bond, J. R., & Gorski, K. M. 2009, ArXiv e-prints
Rocha, G., Contaldi, C. R., Bond, J. R., & Gorski, K. M. 2011, MNRAS, 414, 823R
Rocha, G., Contaldi, C. R., Colombo, L. P. L., et al. 2010, ArXiv e-prints
Stompor, R., Leach, S., Stivoli, F., & Baccigalupi, C. 2009, MNRAS, 392, 216
Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. 2003, Phys.Rev.D, 68, 123523
Tristram, M., Patanchon, G., Macías-Pérez, J. F., et al. 2005, A&A, 436, 785
Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. 2004, Phys. Rev. D, 70, 083511
Wehus et al. 2013, in preparation
Ysard, N., Miville-Deschênes, M. A., & Verstraete, L. 2010, A&A, 509, L1
Ysard, N. & Verstraete, L. 2010, A&A, 509, A12

Appendix A: Physical parametrization

The `Commander-Ruler` (`C-R`) approach implements Bayesian component separation in pixel space, fitting a parametric model to the data by sampling the posterior distribution for the model parameters. For computational reasons, the fit is performed in a two-step procedure: First, both foreground amplitudes and spectral parameters are found at low-resolution using Markov Chain Monte Carlo (MCMC)/Gibbs sampling algorithms (Jewell et al. 2004; Wandelt et al. 2004; Eriksen et al. 2004, 2007, 2008). Second, the amplitudes are recalculated at high resolution by solving the generalized least squares system (GLSS) per pixel with the spectral parameters fixed to the their values from the low-resolution run.

For the CMB-oriented analysis presented in this paper, we only use the seven lowest *Planck* frequencies, i.e., from 30 to 353 GHz. We first downgrade each frequency map from its native angular resolution to a common resolution of 40 arcminutes and repixelize at `HEALPix` $N_{\rm side} = 256$. Second, we set the monopoles and dipoles for each frequency band as described in Sect. 8.1 using a method that locally records spectral indices (Wehus et al. 2013). We approximate the effective instrumental noise as white with a root mean square (RMS) per pixel given by the *Planck* scanning pattern and an amplitude calibrated by smoothing simulations of the instrumental noise including correlations to the same resolution. For the high-resolution analysis, the important pre-processing step is the upgrading of the effective low-resolution mixing matrices to full *Planck* resolution: this is done by repixelizing from $N_{\rm side} = 256$ to 2048 in harmonic space, ensuring that potential pixelization effects from the low-resolution map do not introduce sharp boundaries in the high-resolution map.

Our model for the data, a map $\mathbf{d}_\nu$ at frequency ν, consists of a linear combination of $N_{\rm c}$ astrophysical components, plus instrumental noise,

$$\mathbf{d}_\nu = \sum_{i=1}^{N_{\rm c}} \mathbf{F}_\nu^i(\theta) \cdot \mathbf{A}^i + \mathbf{n}_\nu, \tag{A.1}$$

where $\mathbf{A}^i$ denotes a sky map vector containing the foreground amplitude map for component i normalized at a reference frequency, and $\mathbf{F}_\nu^i(\theta^i)$ is a diagonal matrix describing the spectral emission law for component i as a function of frequency and which depends on a (small) set of spectral parameters, θ. The CMB signal is included in the sum and, as a special case, it may be represented either in harmonic or pixel space, depending on whether the main goal of the analysis is CMB power spectrum analysis or component separation. The former representation is used for the `Commander`-based low-ℓ likelihood presented in Planck Collaboration XV (2013), while the latter is used for the foreground fits presented in Section 8 of this paper.

Bayes theorem specifies the posterior distribution for the model parameters,

$$P(\mathbf{A}^i, \theta|\mathbf{d}) \propto \mathcal{L}(\mathbf{A}^i, \theta) P(\mathbf{A}^i, \theta), \tag{A.2}$$

where $\mathcal{L}(\mathbf{A}^i, \theta)$ is a Gaussian likelihood, and the prior $P(\mathbf{A}^i, \theta^i)$ depends on the application. In this paper, the prior on spectral indices is a product of a Jeffreys prior and physical priors, as detailed in Section 8.2; no priors are imposed on the foreground amplitudes.

In the low-resolution `Commander` analysis, we exploit a Gibbs sampler to map out the posterior distribution (Eriksen et al. 2008), adopting the following minimal two-step scheme:

$$\mathbf{A} \leftarrow P(\mathbf{A}|\theta, \mathbf{d}) \tag{A.3}$$

$$\theta \leftarrow P(\theta|\mathbf{A}, \mathbf{d}). \tag{A.4}$$

The first conditional distribution is a multivariate Gaussian, while the second distribution does not have an analytic form and must be mapped out numerically.

The high-resolution Ruler analysis maximizes the foreground amplitude conditional in Eq. A.3 numerically by solving the generalized least squares system

$$\mathbf{A}^i(\theta) = \Big(\sum_\nu \mathbf{F}_\nu^{i,\mathrm{T}}\mathbf{N}_\nu^{-1}\mathbf{F}_\nu^i\Big)^{-1}\sum_\mu \mathbf{F}_\mu^{i,\mathrm{T}}\mathbf{N}_\mu^{-1}\mathbf{D}_\mu \equiv \sum_\mu \mathbf{W}(\theta)_\mu^i\mathbf{D}_\mu, \tag{A.5}$$

were $\mathbf{N}_\nu$ is the noise covariance matrix (assumed to be diagonal) of the νth channel, $\mathbf{F}^i \equiv \mathbf{F}^i(\theta)$ and we have neglected the different angular resolutions of the channels. The posterior marginal average for the high-resolution amplitude maps is then given by $\langle\mathbf{A}^i\rangle = \int \mathbf{A}^i(\theta)P(\theta)d\theta \simeq \frac{1}{N_{\mathrm{sample}}}\sum_{\theta,\nu}\mathbf{W}(\theta)_\nu^i\mathbf{D}_\nu \equiv \sum_\nu \mathbf{W}_\nu^i\mathbf{D}$, a sum over the N_{sample} samples of the spectral parameters θ.

Once the channel weights, $\mathbf{W}_\nu$, have been computed, processing a large number of simulations requires negligible computational resources. This feature has been extensively used for computation of the effective beam of Ruler maps: FFP6 CMB simulations for the 30 to 353 GHz channels are combined according to $\mathbf{W}_\nu$ and the effective beam transfer function is found as $b_\ell^2 \equiv \langle C_\ell^{\mathrm{out}}/(C_\ell^{\mathrm{inp}}w_\ell^2)\rangle$. Here, C_ℓ^{inp} is the power spectrum of the input simulation before convolution with the instrumental beam, w_ℓ is the HEALPix pixel window function, and the average is taken over the set of simulations. Missing pixels are set to 0 when computing C_ℓ in the above expression. The low number (~ 500) of missing pixels in the data renders the impact of such a choice negligible at $\ell < 2000$. A similar procedure is used for defining the effective beam of the non-CMB components.

The above algorithm produces a set of samples drawn from the posterior distribution, as opposed to a direct estimate of individual component amplitudes or spectral parameters. While this sample set provides a statistically complete representation of the posterior, it is non-trivial to visualize or to compare the distribution with external data. For convenience, we therefore summarize the distribution in terms of mean and standard deviation maps for each component. We emphasize, however, that the distribution is significantly non-Gaussian, and when searching for features in the maps at low signal-to-noise levels, one must take into account the exact distribution.

Finally, the Commander-Ruler confidence mask (see Sect. 5) is primarily defined by the product of the CG80 mask and the point source mask described in Sect. 4. We additionally remove any pixels excluded by the 13% Commander likelihood mask described by Planck Collaboration XV (2013); however, this is almost entirely included within the CG80 mask, and this step therefore has very little impact on the final result. To complete the mask, we remove any pixels for which the high-resolution Ruler CMB map, smoothed to 40 arcminutes, differs by more than $3\,\sigma$ from the low-resolution Commander CMB map, which can happen due to spatial spectral variations on pixel scales.

Appendix B: Internal linear combination

NILC is a method to extract the CMB (or any component with known spectral behaviour) by applying the ILC technique to multi-channel observations in needlet space, that is, with weights that are allowed to vary over the sky and over the full multipole range.

The ability to linearly combine input maps varying over the sky and over multipoles is called localization. In the needlet framework, harmonic localization is achieved using a set of bandpass filters defining a series of scales, and spatial localization is achieved at each scale by defining zones over the sky. The harmonic localization adopted here uses nine spectral bands covering multipoles up to $\ell = 3200$ (see Fig. B.1). The spatial localization depends on the scale. At the coarsest scale, which includes the multipoles of lowest degree, we use a single zone (no localization), while at the finest scales (which include the highest multipoles) the sky is partitioned into 20 zones (again, see Fig. B.1).

The NILC method amounts to computing an ILC in each zone of each scale, allowing the ILC weights to adapt naturally to the varying strength of the other components as a function of position and multipole. A complete description of the basic NILC method can be found in Delabrouille et al. (2009).

In the present work, however, there is an important difference in the processing of the coarsest scale. Since the coarsest scale of the NILC filter is not localized, a plain NILC map would be equivalent to a pixel-based ILC for all the multipoles of that scale. This procedure, however, is known to be quite susceptible to the "ILC bias" due to chance correlations between the CMB and foregrounds. In order to mitigate this effect, the (single) covariance matrix which determines the ILC coefficients at the coarsest scale is not computed as a pixel average, but is rather estimated in the harmonic domain as an average over spherical harmonic coefficients using a spectral weight which equalizes the power of the CMB modes (based on a fiducial spectrum). This can be shown significantly to decrease the large scale errors.

In practice, our NILC processing depends on several implementation choices, as follows:

- Input channels: In this work, the NILC algorithm is applied to all Planck channels from 44 to 857 GHz omitting only the 30 GHz channel.
- Pre-processing of point sources: Identical to the SMICA preprocessing (see Appendix D).
- Masking and inpainting: The NILC CMB map is actually produced in a three-step process. In a first step, the NILC weights are computed from covariance matrices evaluated using a Galactic mask removing about 2 % of the sky (and is apodized at 1°). In a second step, those NILC weights are applied to needlet coefficients computed over the complete sky (except for point source masking/subtraction), yielding a NILC CMB estimate over the full sky (except for the point source mask). In short, the weights are computed over a masked sky but are applied to a full sky (up to point sources). In a final step, the pixels masked due to point source processing are replaced by the values of a constrained Gaussian realization (inpainting).
- Spatial localization: The boundaries of the zones used for spatial localisation (shown at Fig. B.1) are obtained as isolevel curves of a low resolution map of Galactic emission.
- Beam control and transfer function: As in the SMICA processing, the input maps are internally re-beamed to a 5′resolution, so the resulting CMB map is automatically synthesized with an effective Gaussian beam of five arcminutes, according to the unbiased nature of the ILC.
- Using SMICA recalibration: In our current implementation, the NILC solution uses the values determined by SMICA for the CMB spectrum, given in Eq. (D.6).

Appendix C: Template fitting

The original SEVEM algorithm produced clean CMB maps at several frequencies through template fitting, followed by an estimation of the CMB power spectrum from these clean maps us-

Table C.1: Linear coefficients, α_j, and templates used to clean individual frequency maps with SEVEM. The first column lists the templates constructed to produce clean maps. Before subtraction, the maps are smoothed to a common resolution. The (353-143) GHz template is constructed at the resolution of the 44 and 70 GHz channels, in order to clean the 44 and 70 GHz maps, respectively. For the rest of the templates, the first map used to construct the templates was filtered with the beam corresponding to the second map and vice versa. Note that for 100, 143 and 217 GHz channels, we give the coefficients used for the largest of the two regions considered in the cleaning, which covers 97 % of the sky.

Template	44 GHz	70 GHz	100 GHz	143 GHz	217 GHz	353 GHz
30–70	3.65×10^{-1}					
30–44		1.25×10^{-1}	-2.35×10^{-2}	2.14×10^{-2}	-1.03×10^{-1}	
44–70			1.67×10^{-1}	1.23×10^{-1}	1.76×10^{-1}	
217–100						-0.12×10^{1}
217–143						8.99×10^{-1}
353–143	4.05×10^{-3}	9.31×10^{-3}				
545–217						9.92×10^{-2}
545–353			5.21×10^{-3}	7.52×10^{-3}	1.84×10^{-2}	
857–545			-4.66×10^{-5}	-6.67×10^{-5}	-1.21×10^{-4}	-5.02×10^{-4}

ing a method based on the Expectation Maximization algorithm (Martınez-Gonzalez et al. 2003; Leach et al. 2008; Fernández-Cobos et al. 2012). From this power spectrum, a multifrequency Wiener-filtered CMB map was produced. For the present work, only the first step of the method, producing clean CMB maps at different frequencies, is considered. In addition, two of these clean maps are optimally combined to produce a final CMB map.

The templates used for cleaning are internal, i.e., they are constructed from *Planck* data, avoiding the need for external data sets, which usually complicate the analyses and may introduce inconsistencies. In the cleaning process, no assumptions about the foregrounds or noise levels are needed, rendering the technique very robust. The fitting can be done in real or wavelet space (using a fast wavelet adapted to the HEALPix pixelization; Casaponsa et al. 2011) to properly deal with incomplete sky coverage. For expediency, however, we fill in the small number of unobserved pixels at each channel with the mean value of their neighbouring pixels before applying SEVEM.

We construct our templates by subtracting two close *Planck* frequency channel maps, after first smoothing them to a common resolution, and converting to CMB temperature units if necessary, to ensure that the CMB signal is properly removed. A linear combination of the templates, t_j, is then subtracted from (hitherto unused) map, d, to produce a clean CMB map at that frequency. This is done either in real or wavelet space (i.e., scale by scale) at each position on the sky,

$$T_c(\boldsymbol{x}, \nu) = d(\boldsymbol{x}, \nu) - \sum_{j=1}^{n_t} \alpha_j t_j(\boldsymbol{x}), \qquad \text{(C.1)}$$

where n_t is the number of templates. If the cleaning is performed in real space, the α_j coefficients are obtained by minimizing the variance of the clean map, T_c, outside a given mask. When working in wavelet space, the cleaning is performed in the same way at each wavelet scale independently (i.e., the linear coefficients depend on the scale). Although we exclude very contaminated regions during the minimization, the subtraction is performed for all pixels and, therefore, the cleaned maps cover the full-sky (although foreground residuals are expected to be present in the excluded areas).

An additional level of flexibility may also be considered: the linear coefficients can be fixed over the full sky, or in several regions. The regions are then combined in a smooth way, by weighting the pixels at the boundaries to reduce possible discontinuities in the clean maps.

Since the method is linear, we may easily propagate the noise properties to the final CMB map. Moreover, it is very fast and permits the generation of thousands of simulations to characterize the statistical properties of the outputs, a critical need for many cosmological applications. The final CMB map retains the angular resolution of the original frequency map.

There are several configurations possible for SEVEM, depending the number of frequency maps to be cleaned or the number of templates used in the fitting. Note that the production of clean maps at different frequencies is of great interest in order to test the robustness of the results. Therefore, to define the best strategy, one needs to find a compromise between the number of maps that can be cleaned independently and the number of templates that can be constructed.

In particular, we have cleaned the 143 GHz and 217 GHz maps using four templates constructed as the difference of the following *Planck* channels (smoothed to a common resolution): (30−44), (44−70), (545−353) and (857−545). For simplicity, the two maps have been cleaned in real space, since there was no significant improvement when using wavelets, especially at high latitude. In order to take into account the different spectral behaviour of the foregrounds at low and high Galactic latitudes, we considered two independent sky regions, using different sets of coefficients (see Table C.1 for the values of the linear coefficients for the main considered region). The first region corresponds to the brightest 3 % of Galactic emission, while the second region is defined by the remaining 97 % of the sky (see Sect. 4 for a detailed description of our Galactic mask construction). For the first region, the coefficients are actually estimated over the complete sky (we find that this is better than performing the minimization on only the brightest 3 % of the sky, where the CMB is very sub-dominant), while for the second region, we exclude the bright 3 % sky fraction, point sources detected at any frequency and those pixels which have not been observed in all channels.

Note that, for consistency, we have used the same configuration (four templates, cleaning in real space, two regions) for the analysis of the FFP6 simulations. However, we find that this simple configuration produces more contaminated CMB maps than for the data (although the region outside the confidence mask still has low contamination), indicating some differences between the foreground level in the data and in simulations. Therefore, conclusions derived from the FFP6 results for SEVEM should be taken with caution, since they are expected to provide overestimated residuals.

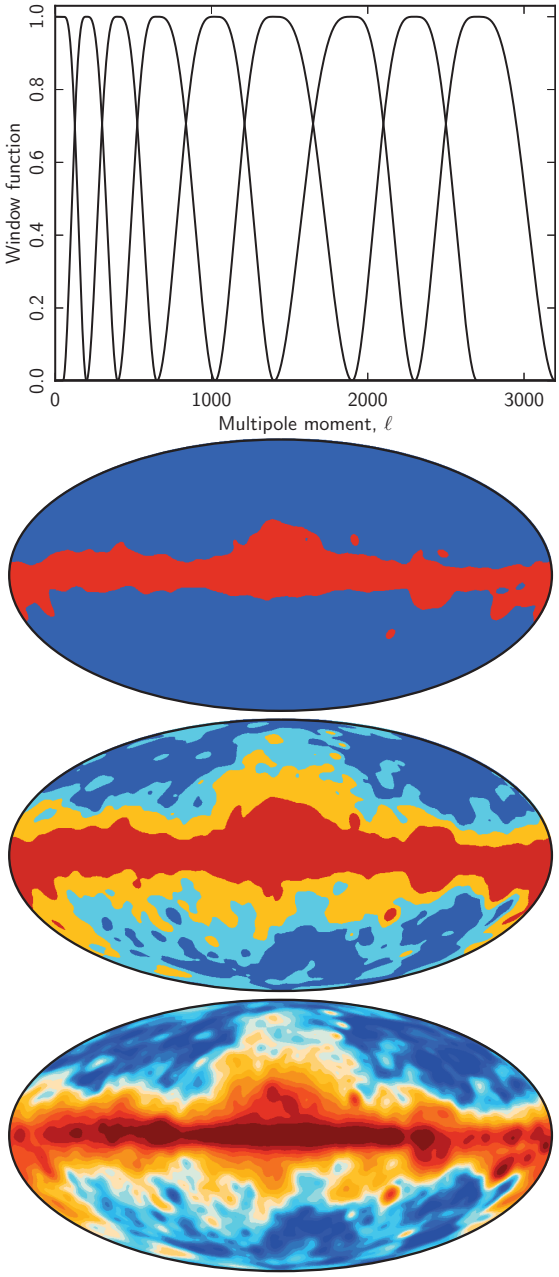


Fig. B.1: Spectral localization for NILC using nine spectral window functions defining nine needlet scales (*top panel*). The scale-dependent spatial localization partitions the sky in one zone (for scale 1), two zones (for scale 2), four zones (for scale 3), or twenty zones (for scales 5, 6, 7, 8, 9). The two-zone, 4-zone and 20-zone partitions are depicted in the lower panels.

Our final CMB map was constructed by combining the 143 and 217 GHz maps by weighting the maps in harmonic space, taking into account the noise level, the resolution, and a rough estimation of the foreground residuals of each map (obtained from realistic simulations). This final map has a resolution corresponding to a Gaussian beam of 5′ FWHM.

Moreover, additional clean CMB maps (at frequencies 44, 70, 100 and 353 GHz) were also produced using different combinations of templates. In particular, to clean the 100 GHz map, we used the same templates and regions as for 143 and 217 GHz. This allows us to produce three (almost) independent clean maps, in the sense that none of the three maps to be cleaned is used to construct the templates. For 44, 70 and 353 GHz, different combinations of templates are used and the linear coefficients are chosen to be the same over the full sky. They are obtained by minimizing the variance of the map outside the same mask as that used to clean the central frequency maps on the largest region. The templates and the corresponding linear coefficients used for each of the considered frequencies are given in Table C.1.

The SEVEM clean frequency maps have been used in analyses of the isotropy and statistics of the CMB (Planck Collaboration XXIII 2013) and to obtain cosmological constraints from the integrated Sachs-Wolfe effect (Planck Collaboration XIX 2013). In particular, clean maps from 44 to 353 GHz were used for the stacking analysis presented in Planck Collaboration XIX (2013), while frequencies from 70 to 217 GHz were used for consistency tests in Planck Collaboration XXIII (2013).

The confidence mask provided for SEVEM is constructed by combining four types of selected regions. In particular, it excludes zones with high residuals identified through the subtraction of SEVEM clean maps at different frequencies, as well as the sources detected at all Planck channels using the Mexican Hat Wavelet algorithm (Planck Collaboration XXVIII 2013). These point sources are masked with holes of varying radius, according to the flux of each source. The pixels that are not observed by all channels are also masked. Finally, to ensure that the area left outside the mask is statistically robust, we also exclude from the analysis the brightest 20 % of Galactic emission, leaving a useful area of around 76 %. This provides a conservative mask for CMB analysis; however, we point out that smaller masks could also be used in specific applications, such as the lensing potential reconstruction described in Sect. 7).

Appendix D: Spectral matching

SMICA (Spectral Matching Independent Component Analysis) reconstructs a CMB map as a linear combination in harmonic space of $N_{\rm chan}$ input frequency maps with weights that depend on multipole ℓ. Given the $N_{\rm chan} \times 1$ vector $\mathbf{x}_{\ell m}$ of spherical harmonic coefficients for the input maps, it computes coefficients $\hat{s}_{\ell m}$ for the CMB map as

$$\hat{s}_{\ell m} = \mathbf{w}_\ell^\dagger \mathbf{x}_{\ell m}, \tag{D.1}$$

where the $N_{\rm chan} \times 1$ vector $\mathbf{w}_\ell$ containing the multipole-dependent weights is chosen to give unit gain to the CMB with minimum variance. This is achieved with

$$\mathbf{w}_\ell = \frac{\mathbf{C}_\ell^{-1}\mathbf{a}}{\mathbf{a}^\dagger \mathbf{C}_\ell^{-1}\mathbf{a}}, \tag{D.2}$$

where vector $\mathbf{a}$ is the spectrum of the CMB evaluated at each channel (allowing for possible inter-channel re-calibration factors) and $\mathbf{C}_\ell$ is the $N_{\rm chan} \times N_{\rm chan}$ spectral covariance matrix of

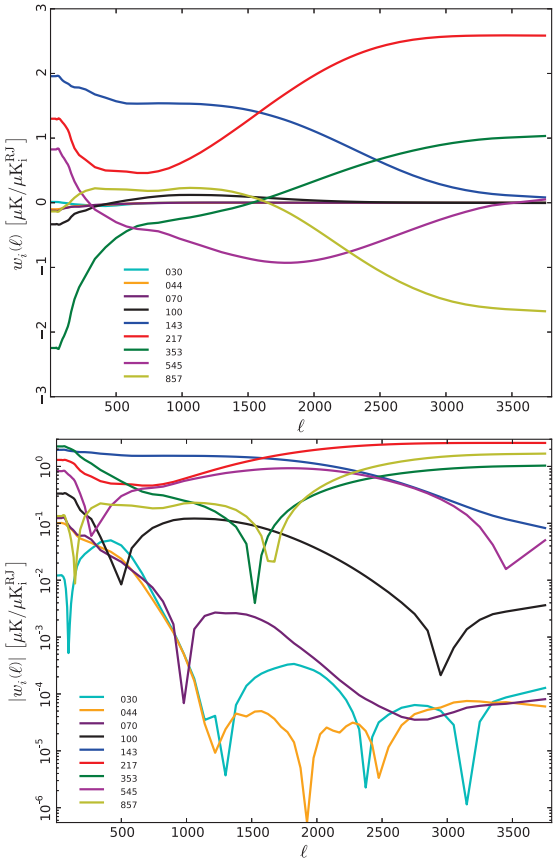


Fig. D.1: Weights, $\mathbf{w}_\ell$, given by SMICA to the input maps, after they are re-beamed to 5′and expressed in $K_{\rm RJ}$, as a function of multipole. Top panel: linear scale; bottom panel: the absolute value of the weights on a logarithmic scale.

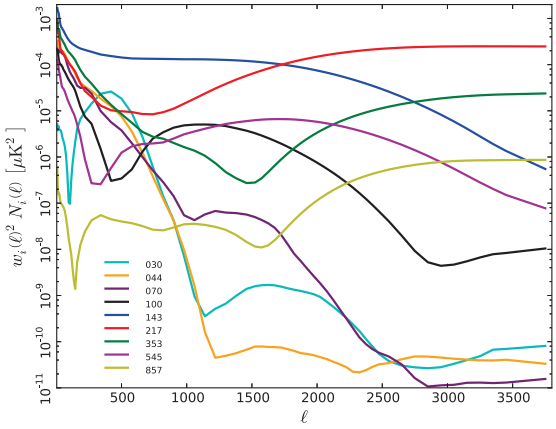


Fig. D.2: Contribution of each input channel to the noise in the SMICA map.

$\mathbf{x}_{\ell m}$. Taking $\mathbf{C}_\ell$ in Eq. (D.2) to be the sample spectral covariance matrix $\widehat{\mathbf{C}}_\ell$ of the observations,

$$\widehat{\mathbf{C}}_\ell = \frac{1}{2\ell+1}\sum_m \mathbf{x}_{\ell m}\mathbf{x}_{\ell m}^\dagger, \tag{D.3}$$

would implement a simple harmonic-domain ILC similar to (Tegmark et al. 2003). At the largest scales, we instead use a model, $\mathbf{C}_\ell(\theta)$, and determine the covariance matrix to be used in Eq. (D.2) by fitting $\mathbf{C}_\ell(\theta)$ to $\widehat{\mathbf{C}}_\ell$. This is done in the maximum likelihood sense for stationary Gaussian fields, that is, the best fit matrices, $\mathbf{C}_\ell(\hat{\theta})$, are obtained for

$$\hat{\theta} = \arg\min_\theta \sum_\ell (2\ell+1)\left(\widehat{\mathbf{C}}_\ell \mathbf{C}_\ell(\theta)^{-1} + \log\det \mathbf{C}_\ell(\theta)\right). \tag{D.4}$$

Equations D.1-D.4 summarize the basic principles of SMICA; its actual operation depends on a choice for the spectral model $\mathbf{C}_\ell(\theta)$, and on several implementation-specific details, which we briefly describe below.

We model the data as a superposition of CMB, noise and foregrounds. The latter are not parametrically modelled; instead, we represent the total foreground emission by d templates with arbitrary frequency spectra, angular spectra and correlations. In the spectral domain, this is equivalent to modelling the covariance matrices as

$$\mathbf{C}_\ell(\theta) = \mathbf{a}\mathbf{a}^\dagger C_\ell + \mathbf{A}\mathbf{P}_\ell\mathbf{A}^\dagger + \mathbf{N}_\ell, \tag{D.5}$$

where C_ℓ is the angular power spectrum of the CMB, $\mathbf{A}$ is a $N_{\rm chan} \times d$ matrix, $\mathbf{P}_\ell$ is a positive definite $d \times d$ matrix, and $\mathbf{N}_\ell$ is a diagonal matrix representing the noise power spectra of the data. The parameter vector θ contains all or part of the quantities in (D.5).

The decomposition D.5 reflects the fact that CMB, foregrounds and noise are independent components of the signal. Thus, SMICA is an ICA (independent component analysis) method. It operates by matching the observations $\widehat{\mathbf{C}}$ to the spectral model (D.5) using the criterion (D.4).

The maximal flexibility in a SMICA fit of model (D.5) is obtained with all the parameters free, that is without any constraint on the spectrum C_ℓ, on the diagonal entries of $\mathbf{N}_\ell$, on $\mathbf{a}$, or on $\mathbf{A}$ and $\mathbf{P}_\ell$. One would ideally fit all those parameters (except for obvious degeneracies, like that between a scale factor in $\mathbf{a}$ and the overall normalization of the CMB spectrum C_ℓ) over the whole multipole range. In practice, this turns out to be too difficult given the large dynamic range both over the sky and over multipoles. We resort to a pragmatic three-step approach in which the criterion (D.4) is minimized by first fitting $\mathbf{a}$, then $\mathbf{A}$, and finally the linear parameters C_ℓ and $\mathbf{N}_\ell$. Each fit is conducted over the multipole ranges and the sky fraction most appropriate for the parameter of interest, as follows.

We first estimate the CMB spectral law $\mathbf{a}$ by fitting all model parameters (that is, without constraint) over a clean fraction of sky (f_{sky}=40 %) in the range $100 \le \ell \le 680$ where the signal is CMB-dominated in most of the channels and the beam window functions are accurately known. In this fit, which is done over a clean part of the sky, we use a foreground emission matrix, $\mathbf{A}$, with only four columns. From this step, we only retain the best fit value for vector $\mathbf{a}$. In the second step, we estimate the foreground emissivity by fixing $\mathbf{a}$ to its value from the previous

step and fitting all the other parameters over a large fraction of sky (f_{sky}=97 %) in the range $4 \le \ell \le 150$ where the signal is dominated by the Galactic emission in all channels. From this second step, we retain the best fit value for the matrix $\mathbf{A}$ which, again, is adjusted without constraint other than having $d = 6$ columns. In the last step, we fit all power spectrum parameters: we fix $\mathbf{a}$ and $\mathbf{A}$ to their previously found values and fit for C_ℓ and $\mathbf{P}_\ell$ at each ℓ.

Note that the first step (fitting $\mathbf{a}$) amounts to re-calibrating the input maps on the basis of CMB anisotropies. For the maps in thermodynamics units, we find

$$\hat{\mathbf{a}} = [0.9900, 1.0000, 1.0020, 0.9990, 1.0000, 1.0004, 0.9920, 1.0457, 1.0000] \quad \text{(D.6)}$$

The value at 857 GHz is not accurately recovered by SMICA, so we have set $a_{857} = 1$. Since the norm of $\mathbf{a}$ is degenerate with a global scale factor for the CMB angular spectrum, it can only be recovered by SMICA up to a scale factor. This degeneracy is fixed here by taking $a_{143} = 1$. The re-calibration step could have been omitted since $\hat{\mathbf{a}}$ is very close the unit vector. However, we found that using $\hat{\mathbf{a}}$ improved the behavior of SMICA over using $\mathbf{a} = [1, \ldots, 1]$.

Before describing implementation details, we explain how SMICA deals with the varying resolution of the input channels, since the discussion thus far assumed that all input maps had the same resolution. Since SMICA works in the harmonic domain, it is a simple matter to account for the beam transfer function, $b_i(\ell)$, of the i-th input map. The CMB sky multipoles $s_{\ell m}$ contribute $s_{\ell m} a_i b_i(\ell) p_i(\ell)$ to the harmonic coefficient $x^i_{\ell m}$ of the i-th map (where $p_i(\ell)$ is the pixel window function for the HEALPix map at $N^i_{\rm side}$). Therefore, in order to produce a final CMB map at 5′resolution, close to the highest resolution of *Planck*, we only need to work with input spherical harmonics re-beamed to 5′; that is, to apply SMICA on vectors $\tilde{\mathbf{x}}_{\ell m}$ with entries $\tilde{x}^i_{\ell m} = x^i_{\ell m} b_5(\ell)/b_i(\ell)/p_i(\ell)$, where $b_5(\ell)$ is a five-arcminute Gaussian beam function. By construction, SMICA then produces an CMB map with an effective Gaussian beam of 5′ (without the pixel window function).

We now give further details on the actual implementation of SMICA:

- Inputs: SMICA uses all nine *Planck* frequency channels from 30 to 857 GHz, harmonically transformed up to $\ell = 4000$.
- Pre-processing of point sources: SMICA is applied on input maps in which point sources are subtracted or masked. We start by fitting the PCCS point sources with SNR > 5 to a Gaussian shape where the source amplitude is estimated together with its position and a constant factor representing the background variance. If the fit is successful ($\chi^2 \le 2$), the fitted point source is removed from the map; otherwise it is masked in all channels and the hole is inpainted by a simple diffusive filling process. This is done at all frequencies except 545 and 857 GHz, where all point sources with SNR > 7.5 are masked and inpainted.
- Beams: When the harmonic coefficients of the input maps are re-beamed at 5′, we do not apply exactly the expression $\tilde{x}^i_{\ell m} = x^i_{\ell m} b_5(\ell)/b_i(\ell)/p_i(\ell)$ mentioned above because the factor $1/b_i(\ell)$ would diverge at high ℓ for the lowest resolution input channels. That may not be a problem in infinite precision arithmetic, but would lead to matrices $\widehat{\mathbf{C}}(\ell)$ with extremely large condition numbers. Instead, we re-beam with the factor $1/b_i(\ell)$ replaced by $\min(1/b_i(\ell), 1000)$. The re-beaming of the CMB modes then is no longer perfect, but this is of course irrelevant because the thresholding occurs in a regime where the signal is completely dominated by the noise, so that the contribution of the corresponding channel is already highly attenuated by the SMICA weights (as shown in Fig. D.1).
- Masking: In practice, SMICA operates on a masked sky, the mask being applied after the point source processing. The mask is obtained by thresholding a heavily smoothed version of the point source mask. The threshold is chosen to leave about 97 % of the sky. Because of the heavy smoothing, the mask has smooth contours and is only sensitive to large aggregates of point sources: the masked areas mostly lie in the Galactic plane, but include also a few bright regions like the Large Magellanic Cloud.
- Inpainting: The SMICA map used in this paper has no real power in the masked region described above. However, for convenience, an inpainted SMICA map has also been produced by replacing the masked pixels with a constrained Gaussian realization obtained by the method of Benoit-Lévy et al. (2013). That map appears in Planck Collaboration I (2013).
- Binning: In our implementation, we use binned spectra.
- Processing at fine scales: Since there is little point trying to model the spectral covariance at high multipoles, because the sample estimate is sufficient, SMICA implements a simple harmonic ILC at $\ell > 1500$; that is, it applies the filter (D.2) with $\mathbf{C}_\ell = \widehat{\mathbf{C}}_\ell$.
- Confidence mask: A confidence mask (Fig. 3) is provided with SMICA, constructed in the following way. The SMICA CMB map is bandpass filtered through a spectral window $v(\ell) = \exp[-((\ell - 1700)/200)^2/2]$. The result is squared and smoothed at two-degree resolution, yielding a map of the (bandpassed) variance of the CMB map. That variance is corrected for the noise contribution by subtracting the variance map for the noise obtained by the same procedure applied to the SMICA HRHD map. If the SMICA map contained only CMB and noise, the variance map would have a uniform value $\sum_\ell v(\ell)^2 b_5(\ell)^2 C(\ell)(2\ell + 1)/4\pi = 31.3\,\mu\mathrm{K}^2$ over the sky. The confidence map is obtained by thresholding the noise-corrected variance map at $70\,\mu\mathrm{K}^2$.

Viewed as a filter, SMICA can be summarized by the weights $\mathbf{w}_\ell$ applied to each input map as a function of multipole. In this sense, SMICA is strictly equivalent to co-adding the input maps after convolution by specific axisymmetric kernels directly related to the corresponding entry of $\mathbf{w}_\ell$.

The SMICA weights used here are shown in Figure D.1 (for input maps in units of $K_{\rm RJ}$). We see, in particular, the (expected) progressive attenuation of the lowest resolution channels with increasing multipole. Figure D.2 shows the contribution of each input channel to the noise in the SMICA CMB map as a function of multipole. The spectral noise contribution from channel i is simply obtained as $w(\ell)^2 N_i(\ell)$, where $w_i(\ell)$ is the i-th entry of the weight vector $\mathbf{w}(\ell)$ and $N_i(\ell)$ is the angular spectrum of the i-th noise map.

More details about SMICA are given in Cardoso et al. (2008), as well as in applications to the analysis of *WMAP* (Patanchon et al. 2005) and Archeops data (Tristram et al. 2005). An application to the measurement of the tensor to scalar ratio using CMB B-modes is discussed in Betoule et al. (2009). Within the Planck collaboration, SMICA is used to define the Plik high-ℓ likelihood (Planck Collaboration XV 2013), but physical models of foreground emission are used there instead of the non-parametric foreground model used here. SMICA is also used to

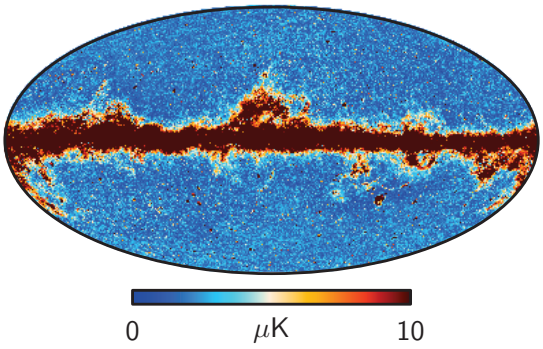


Fig. E.1: Standard deviation of the four foreground-cleaned CMB maps from the FFP6 simulation. All maps have been downgraded to a `HEALPix` resolution of $N_{\rm side} = 128$.

cross-check the HFI calibration (see Planck Collaboration VIII (2013)).

Appendix E: FFP6 simulation results

We study the performance of the component separation algorithms on the FFP6 simulation described in Sect. 4, providing additional information beyond that in the body of the paper. Much of the analysis presented here mirrors that shown for the data in Sects. 5 and 6, allowing a direct comparison between the two analyses.

E.1. CMB maps

First, we show in Fig. E.1 the standard deviation between the four foreground-cleaned FFP6 maps, similar to that shown in Fig. 5 for the data. Figure E.2 shows all pairwise differences between the same maps, mirroring Fig. 6 for the data. These two plots highlight an important point concerning the FFP6 analysis already mentioned in Sect. 5, namely that in near-Galactic regions, where the foregrounds are important, the internal differences between the four algorithms are larger in the FFP6 simulation than in the real data. This is due to the fact that each component separation algorithm has been optimized with the real data in terms of model definition, localization, etc. Then, the same models have been used for the FFP6 simulation without change. Only the parameters within those models are refitted to the new data set. This implies in fact that we expect each method to perform better on the data than the simulations in terms of absolute residuals, to the extent that the simulation matches the real sky. In other words, the FFP6 simulation provides a conservative estimate of the residual errors in the real data.

E.2. Power spectrum residuals from individual components

In Fig. E.3 we show the residual effect of some of the individual components on the foreground-cleaned CMB map. The thermal dust emission, CIB fluctuations, point sources, and noise have been processed individually with each algorithm. All other components (free-free, synchrotron, spinning dust, CO, thermal SZ, and kinetic SZ) are shown as a single, composite residual component.

E.3. CMB power spectra and cosmological parameters

We assess the performance of our component separation techniques by evaluating cosmological constraints from the foreground-cleaned CMB maps derived from the FFP6 simulation.

Figure E.4 shows the estimates of the angular power spectra of the CMB maps. Figure E.5 compares the cosmological parameters derived from the four foreground-cleaned CMB maps, `CamSpec`[6], and `Plik` to the input (theoretical) parameters for different ℓ-ranges. The parameter space is defined by the same model applied to the real data in Sect. 6, including six ΛCDM and two foreground parameters. All deviations from input parameters are small and within $1\,\sigma$ up to $\ell = 2000$, verifying that all methods work well in this multipole range. However, for $\ell_{\rm max} = 2500$ we start to see significant shifts, e.g., for $\Omega_{\rm b}h^2$ and $n_{\rm s}$. Further, the point source foreground parameter, $A_{\rm ps}$, reaches large values, implying that assumptions concerning the high-ℓ foreground model become important. For these reasons, we consider $\ell_{\rm max} = 2000$ as the maximum recommended ℓ-range for these maps in the current data release.

Still, the overall agreement is excellent between all codes and all ℓ-ranges. In particular, we see that differences in the band power spectra at high ℓ between the different codes are mostly absorbed by the two-parameter foreground model. For instance, the `Commander-Ruler` band power spectrum has more power at high ℓ due to noise or residual point sources, but this excess is well fitted by the two-parameter foreground model, and mostly interpreted in terms of a residual point source component; this is expected, given the lower angular resolution of this map.

As mentioned above, n_s and and A are to some extent sensitive to $\ell_{\rm max}$. These parameters are degenerate with the foreground parameters. This may suggest that our C_ℓ foreground templates deviate more from the shape of the Poissonian and clustered component in the CMB map. This is a limitation of the simple foreground templates used here. To properly describe the foreground residuals in the reconstructed maps, we should use a foreground power spectrum template tailored to each method. For instance, such templates may be constructed by processing simulated foreground maps though each of the four pipelines. The templates are then given by the pseudo-C_ℓ of each of the processed foreground map. However, our analysis shows that the current simple model provides accurate results when restricting the analysis to $\ell_{\rm max} = 2000$.

Figure E.6 shows the best-fit power spectrum residuals for the CMB map, `CamSpec` and `Plik` relative to the input CMB ΛCDM model estimated up to $\ell = 2000$. These plots show that the residuals of the CMB map-based best-fit models are comparable to the `CamSpec` and `Plik` residuals, and smaller than $40\,\mu{\rm K}^2$ for most of the ℓ range with larger deviations observed for `CamSpec` at $\ell \sim 200$. At higher ℓs the residuals are smaller than $10\,\mu{\rm K}^2$ for both approaches, all showing similar trends. Thus, both the map- and spectrum-based likelihoods recover input parameters reasonably well, with the latter yielding slightly larger deviations from the best-fit model of the input CMB realization.

[1] APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/lrfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France

[6] For `CamSpec`, $k_{\rm pivot} = 0.05$ was adopted for this test, while all others, input parameters and input CMB realization included, use $k_{\rm pivot} = 0.002$.

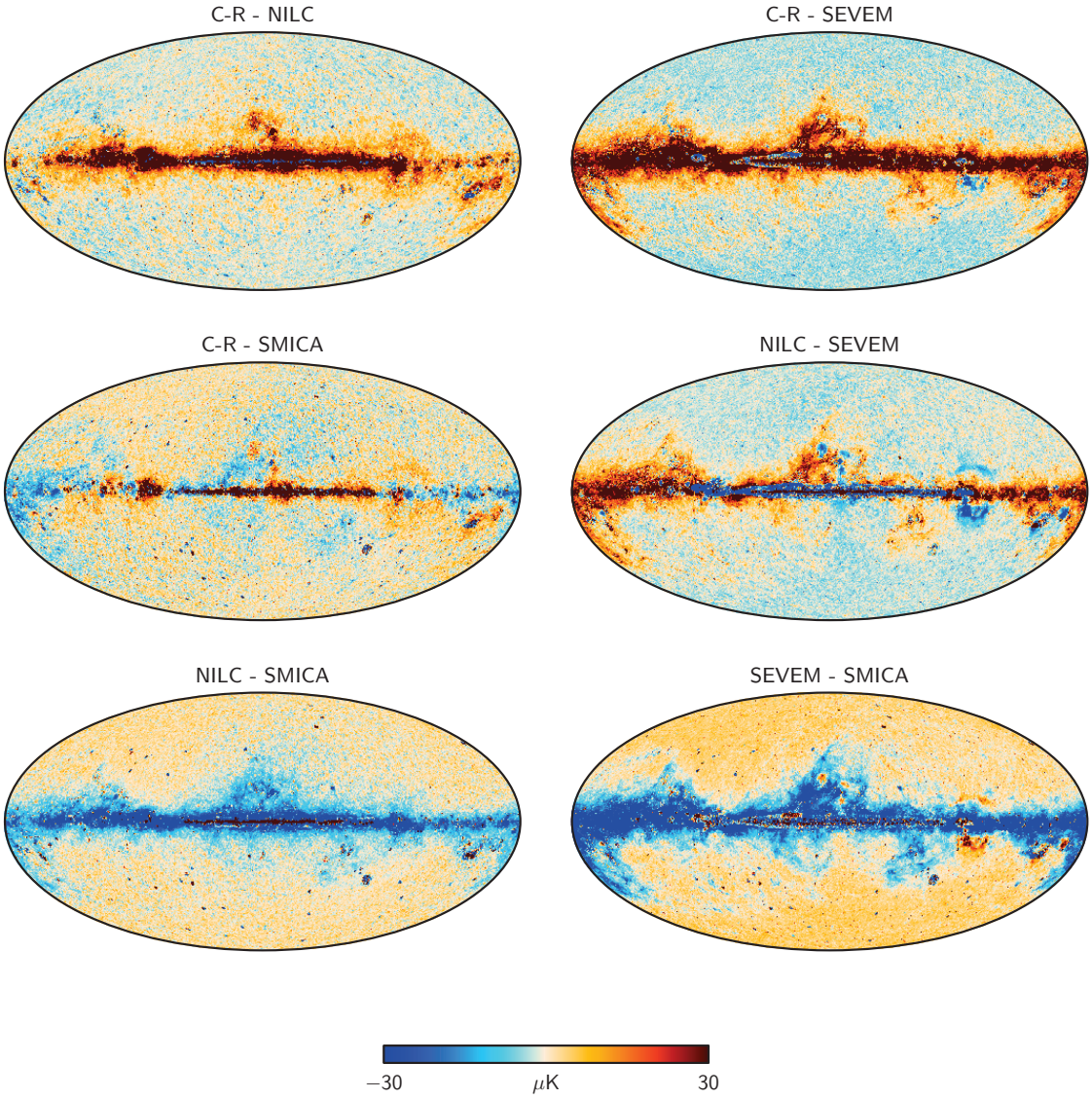


Fig. E.2: Pairwise differences between foreground-cleaned CMB maps from the FFP6 simulation. All maps have been downgraded to a `HEALPix` resolution of $N_{\rm side} = 128$ to show the large-scale differences.

[2] Aalto University Metsähovi Radio Observatory, Metsähovintie 114, FIN-02540 Kylmälä, Finland
[3] African Institute for Mathematical Sciences, 6-8 Melrose Road, Muizenberg, Cape Town, South Africa
[4] Agenzia Spaziale Italiana Science Data Center, c/o ESRIN, via Galileo Galilei, Frascati, Italy
[5] Agenzia Spaziale Italiana, Viale Liegi 26, Roma, Italy
[6] Astrophysics Group, Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, U.K.
[7] Astrophysics & Cosmology Research Unit, School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Private Bag X54001, Durban 4000, South Africa
[8] Atacama Large Millimeter/submillimeter Array, ALMA Santiago Central Offices, Alonso de Cordova 3107, Vitacura, Casilla 763 0355, Santiago, Chile
[9] CITA, University of Toronto, 60 St. George St., Toronto, ON M5S 3H8, Canada
[10] CNR - ISTI, Area della Ricerca, via G. Moruzzi 1, Pisa, Italy
[11] CNRS, IRAP, 9 Av. colonel Roche, BP 44346, F-31028 Toulouse cedex 4, France
[12] California Institute of Technology, Pasadena, California, U.S.A.

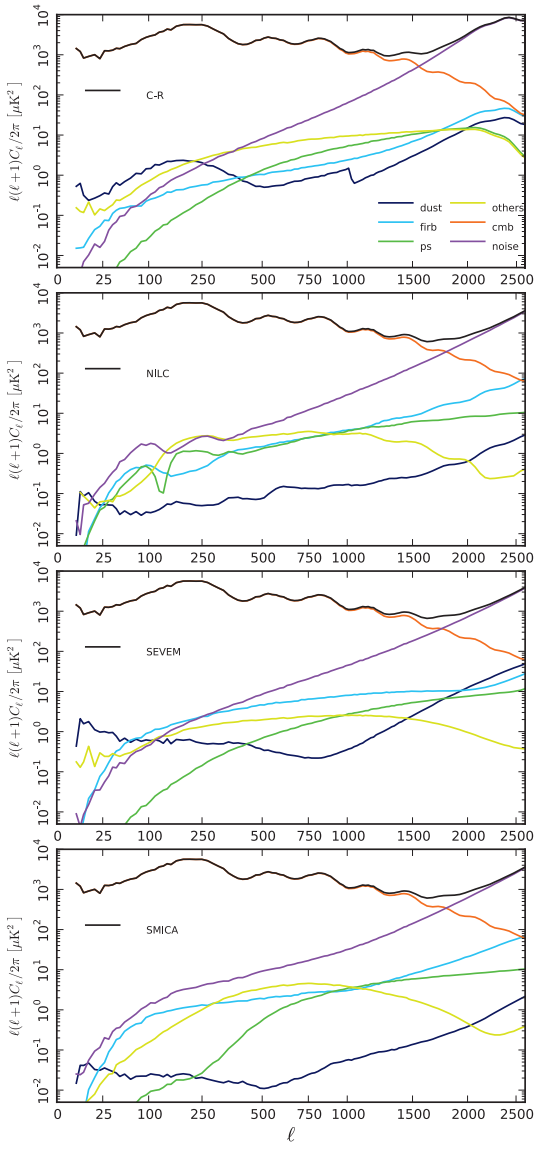


Fig. E.3: Angular power spectra of residual foreground emission in the CMB maps from the FFP6 simulation. The components shown are: thermal dust, cosmic infrared background fluctuations, point sources, CMB, noise, and the sum of all others. From top to bottom, the panels show the results for `Commander-Ruler`, `NILC`, `SEVEM`, and `SMICA`.

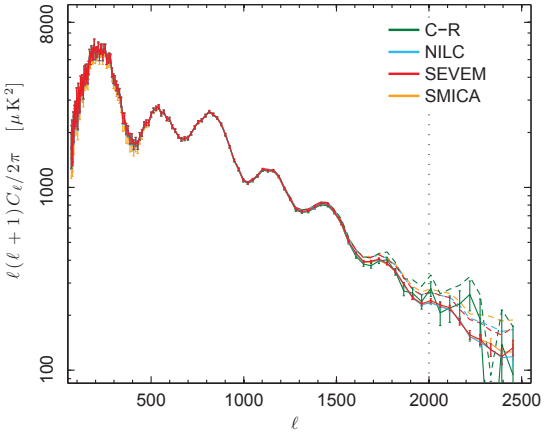


Fig. E.4: Estimates of the CMB power spectra from the foreground-cleaned FFP6 maps, computed by `XFaster`. The solid lines show the spectra after subtracting the best-fit model of residual foregrounds. The vertical dotted line shows the maximum multipole ($\ell = 2000$) used in the likelihood for fitting the foreground model and cosmological parameters (see Sect. E.3 for further details). The dashed lines show the spectra before residual foreground subtraction.

[13] Centre for Theoretical Cosmology, DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA U.K.
[14] Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Plaza San Juan, 1, planta 2, E-44001, Teruel, Spain
[15] Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.
[16] Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain
[17] DSM/Irfu/SPP, CEA-Saclay, F-91191 Gif-sur-Yvette Cedex, France
[18] DTU Space, National Space Institute, Technical University of Denmark, Elektrovej 327, DK-2800 Kgs. Lyngby, Denmark
[19] Département de Physique Théorique, Université de Genève, 24, Quai E. Ansermet,1211 Genève 4, Switzerland
[20] Departamento de Física Fundamental, Facultad de Ciencias, Universidad de Salamanca, 37008 Salamanca, Spain
[21] Departamento de Física, Universidad de Oviedo, Avda. Calvo Sotelo s/n, Oviedo, Spain
[22] Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Avda. de los Castros s/n, Santander, Spain
[23] Department of Astronomy and Astrophysics, University of Toronto, 50 Saint George Street, Toronto, Ontario, Canada
[24] Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
[25] Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California, U.S.A.
[26] Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia, Canada
[27] Department of Physics and Astronomy, Dana and David Dornsife College of Letter, Arts and Sciences, University of Southern California, Los Angeles, CA 90089, U.S.A.
[28] Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.
[29] Department of Physics, Gustaf Hällströmin katu 2a, University of Helsinki, Helsinki, Finland
[30] Department of Physics, Princeton University, Princeton, New Jersey, U.S.A.
[31] Department of Physics, University of California, One Shields Avenue, Davis, California, U.S.A.

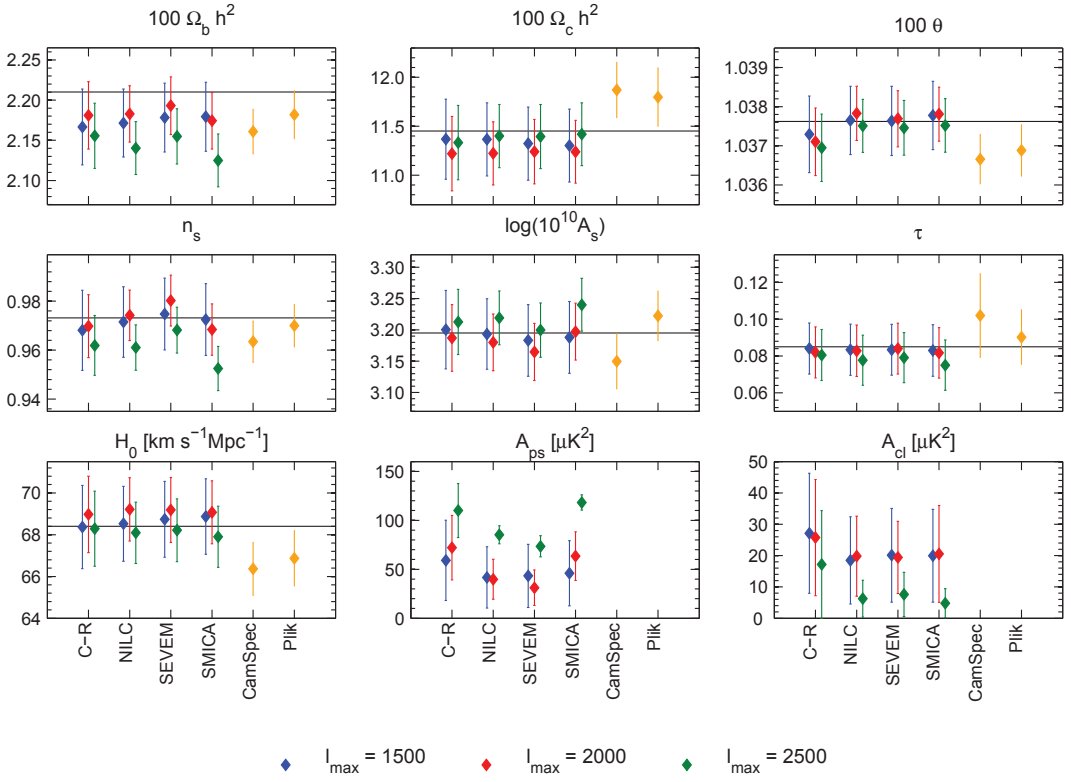


Fig. E.5: Comparison of cosmological parameters derived from the FFP6 simulation using different methods. The parameters shown as blue, red and green points indicate results obtained with $\ell_{\rm max}$ = 1500, 2000 and 2500, respectively, and the yellow points show the results derived by `CamSpec` and `Plik` using cross-spectra. The black horizontal lines mark the input parameter values. The values of the foreground parameters are not shown for `CamSpec` or `Plik` since they use a different model. The matter power spectrum pivot scale was $k_{\rm pivot}$ = 0.002 for all likelihoods, except `CamSpec` for which $k_{\rm pivot}$ = 0.05 was used.

[32] Department of Physics, University of California, Santa Barbara, California, U.S.A.
[33] Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois, U.S.A.
[34] Dipartimento di Fisica e Astronomia G. Galilei, Università degli Studi di Padova, via Marzolo 8, 35131 Padova, Italy
[35] Dipartimento di Fisica e Scienze della Terra, Università di Ferrara, Via Saragat 1, 44122 Ferrara, Italy
[36] Dipartimento di Fisica, Università La Sapienza, P. le A. Moro 2, Roma, Italy
[37] Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria, 16, Milano, Italy
[38] Dipartimento di Fisica, Università degli Studi di Trieste, via A. Valerio 2, Trieste, Italy
[39] Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy
[40] Discovery Center, Niels Bohr Institute, Blegdamsvej 17, Copenhagen, Denmark
[41] Dpto. Astrofísica, Universidad de La Laguna (ULL), E-38206 La Laguna, Tenerife, Spain
[42] European Southern Observatory, ESO Vitacura, Alonso de Cordova 3107, Vitacura, Casilla 19001, Santiago, Chile
[43] European Space Agency, ESAC, Planck Science Office, Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, Madrid, Spain
[44] European Space Agency, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands
[45] Haverford College Astronomy Department, 370 Lancaster Avenue, Haverford, Pennsylvania, U.S.A.
[46] Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland
[47] INAF - Osservatorio Astrofisico di Catania, Via S. Sofia 78, Catania, Italy
[48] INAF - Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, Padova, Italy
[49] INAF - Osservatorio Astronomico di Roma, via di Frascati 33, Monte Porzio Catone, Italy
[50] INAF - Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, Trieste, Italy
[51] INAF/IASF Bologna, Via Gobetti 101, Bologna, Italy
[52] INAF/IASF Milano, Via E. Bassini 15, Milano, Italy
[53] INFN, Sezione di Bologna, Via Irnerio 46, I-40126, Bologna, Italy
[54] INFN, Sezione di Roma 1, Università di Roma Sapienza, Piazzale Aldo Moro 2, 00185, Roma, Italy

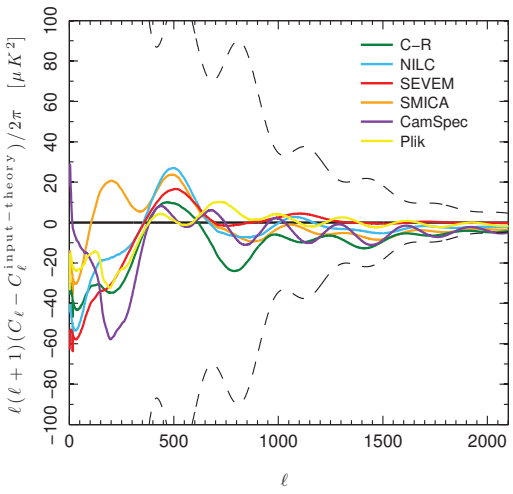


Fig. E.6: Residuals of map-based and spectrum-based best-fit models relative to the FFP6 simulation input ΛCDM spectrum, shown for each algorithm up to $\ell_{\rm max} = 2000$. Cosmic variance is shown as the black dashed line.

[55] INFN/National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy
[56] IPAG: Institut de Planétologie et d'Astrophysique de Grenoble, Université Joseph Fourier, Grenoble 1 / CNRS-INSU, UMR 5274, Grenoble, F-38041, France
[57] ISDC Data Centre for Astrophysics, University of Geneva, ch. d'Ecogia 16, Versoix, Switzerland
[58] IUCAA, Post Bag 4, Ganeshkhind, Pune University Campus, Pune 411 007, India
[59] Imperial College London, Astrophysics group, Blackett Laboratory, Prince Consort Road, London, SW7 2AZ, U.K.
[60] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, U.S.A.
[61] Institut Néel, CNRS, Université Joseph Fourier Grenoble I, 25 rue des Martyrs, Grenoble, France
[62] Institut Universitaire de France, 103, bd Saint-Michel, 75005, Paris, France
[63] Institut d'Astrophysique Spatiale, CNRS (UMR8617) Université Paris-Sud 11, Bâtiment 121, Orsay, France
[64] Institut d'Astrophysique de Paris, CNRS (UMR7095), 98 bis Boulevard Arago, F-75014, Paris, France
[65] Institute for Space Sciences, Bucharest-Magurale, Romania
[66] Institute of Astronomy and Astrophysics, Academia Sinica, Taipei, Taiwan
[67] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, U.K.
[68] Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway
[69] Instituto de Astrofísica de Canarias, C/Vía Láctea s/n, La Laguna, Tenerife, Spain
[70] Instituto de Física de Cantabria (CSIC-Universidad de Cantabria), Avda. de los Castros s/n, Santander, Spain
[71] Istituto di Fisica del Plasma, CNR-ENEA-EURATOM Association, Via R. Cozzi 53, Milano, Italy
[72] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California, U.S.A.
[73] Jodrell Bank Centre for Astrophysics, Alan Turing Building, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester, M13 9PL, U.K.
[74] Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 0HA, U.K.
[75] LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France
[76] LERMA, CNRS, Observatoire de Paris, 61 Avenue de l'Observatoire, Paris, France
[77] Laboratoire AIM, IRFU/Service d'Astrophysique - CEA/DSM - CNRS - Université Paris Diderot, Bât. 709, CEA-Saclay, F-91191 Gif-sur-Yvette Cedex, France
[78] Laboratoire Traitement et Communication de l'Information, CNRS (UMR 5141) and Télécom ParisTech, 46 rue Barrault F-75634 Paris Cedex 13, France
[79] Laboratoire de Physique Subatomique et de Cosmologie, Université Joseph Fourier Grenoble I, CNRS/IN2P3, Institut National Polytechnique de Grenoble, 53 rue des Martyrs, 38026 Grenoble cedex, France
[80] Laboratoire de Physique Théorique, Université Paris-Sud 11 & CNRS, Bâtiment 210, 91405 Orsay, France
[81] Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.
[82] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
[83] McGill Physics, Ernest Rutherford Physics Building, McGill University, 3600 rue University, Montréal, QC, H3A 2T8, Canada
[84] MilliLab, VTT Technical Research Centre of Finland, Tietotie 3, Espoo, Finland
[85] Niels Bohr Institute, Blegdamsvej 17, Copenhagen, Denmark
[86] Observational Cosmology, Mail Stop 367-17, California Institute of Technology, Pasadena, CA, 91125, U.S.A.
[87] Optical Science Laboratory, University College London, Gower Street, London, U.K.
[88] SB-ITP-LPPC, EPFL, CH-1015, Lausanne, Switzerland
[89] SISSA, Astrophysics Sector, via Bonomea 265, 34136, Trieste, Italy
[90] School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff, CF24 3AA, U.K.
[91] School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, U.K.
[92] Space Research Institute (IKI), Russian Academy of Sciences, Profsoyuznaya Str, 84/32, Moscow, 117997, Russia
[93] Space Sciences Laboratory, University of California, Berkeley, California, U.S.A.
[94] Stanford University, Dept of Physics, Varian Physics Bldg, 382 Via Pueblo Mall, Stanford, California, U.S.A.
[95] Sub-Department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, U.K.
[96] Theory Division, PH-TH, CERN, CH-1211, Geneva 23, Switzerland
[97] UPMC Univ Paris 06, UMR7095, 98 bis Boulevard Arago, F-75014, Paris, France
[98] Université de Toulouse, UPS-OMP, IRAP, F-31028 Toulouse cedex 4, France
[99] Universities Space Research Association, Stratospheric Observatory for Infrared Astronomy, MS 232-11, Moffett Field, CA 94035, U.S.A.
[100] University of Granada, Departamento de Física Teórica y del Cosmos, Facultad de Ciencias, Granada, Spain
[101] University of Miami, Knight Physics Building, 1320 Campo Sano Dr., Coral Gables, Florida, U.S.A.
[102] Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland

Paper III

***Planck* 2013 results. XXIII. Isotropy and statistics of the CMB**

Planck Collaboration

Submitted to Astronomy & Astrophysics

Astronomy & Astrophysics manuscript no. IandS˙v6
March 22, 2013

Planck 2013 results. XXIII. Isotropy and statistics of the CMB

Planck Collaboration: P. A. R. Ade[87], N. Aghanim[60], C. Armitage-Caplan[92], M. Arnaud[73], M. Ashdown[70,6], F. Atrio-Barandela[18], J. Aumont[60], C. Baccigalupi[86], A. J. Banday[95,9], R. B. Barreiro[67], J. G. Bartlett[1,68], N. Bartolo[34], E. Battaner[96], R. Battye[69], K. Benabed[61,94], A. Benoît[58], A. Benoit-Lévy[26,61,94], J.-P. Bernard[9], M. Bersanelli[37,50], P. Bielewicz[95,9,86], J. Bobin[73], J. J. Bock[68,10], A. Bonaldi[69], L. Bonavera[67], J. R. Bond[8], J. Borrill[13,89], F. R. Bouchet[61,94], M. Bridges[70,6,64], M. Bucher[1], C. Burigana[49,35], R. C. Butler[49], J.-F. Cardoso[74,1,61], A. Catalano[75,72], A. Challinor[64,70,11], A. Chamballu[73,15,60], R.-R. Chary[57], L.-Y Chiang[63], H. C. Chiang[28,7], P. R. Christensen[82,41], S. Church[91], D. L. Clements[56], S. Colombi[61,94], L. P. L. Colombo[25,68], F. Couchot[71], A. Coulais[72], B. P. Crill[68,83], M. Cruz[20], A. Curto[6,67], F. Cuttaia[49], L. Danese[86], R. D. Davies[69], R. J. Davis[69], P. de Bernardis[36], A. de Rosa[49], G. de Zotti[46,86], J. Delabrouille[1], J.-M. Delouis[61,94], F.-X. Désert[53], J. M. Diego[67], H. Dole[60,59], S. Donzelli[50], O. Doré[68,10], M. Douspis[60], A. Ducout[61], X. Dupac[43], G. Efstathiou[64], F. Elsner[61,94], T. A. Enßlin[78], H. K. Eriksen[65], Y. Fantaye[65], J. Fergusson[11], F. Finelli[49,51], O. Forni[95,9], M. Frailis[48], E. Franceschi[49], M. Frommert[17], S. Galeotta[48], K. Ganga[1], M. Giard[95,9], G. Giardino[44], Y. Giraud-Héraud[1], J. González-Nuevo[67,86], K. M. Górski[68,98], S. Gratton[70,64], A. Gregorio[38,48], A. Gruppuso[49], M. Hansen[82], F. K. Hansen[65], D. Hanson[80,68,8], D. Harrison[64,70], G. Helou[10], S. Henrot-Versillé[71], C. Hernández-Monteagudo[12,78], D. Herranz[67], S. R. Hildebrandt[10], E. Hivon[61,94], M. Hobson[6], W. A. Holmes[68], A. Hornstrup[16], W. Hovest[78], K. M. Huffenberger[97], T. R. Jaffe[95,9], A. H. Jaffe[56], W. C. Jones[28], M. Juvela[27], E. Keihänen[27], R. Keskitalo[23,13], J. Kim[82], T. S. Kisner[77], J. Knoche[78], L. Knox[31], M. Kunz[17,60,3], H. Kurki-Suonio[27,45], G. Lagache[60], A. Lähteenmäki[2,45], J.-M. Lamarre[72], A. Lasenby[6,70], R. J. Laureijs[44], C. R. Lawrence[68], J. P. Leahy[69], R. Leonardi[43], C. Leroy[60,95,9], J. Lesgourgues[93,85], M. Liguori[34], P. B. Lilje[65], M. Linden-Vørnle[16], M. López-Caniego[67], P. M. Lubin[32], J. F. Macías-Pérez[75], B. Maffei[69], D. Maino[37,50], N. Mandolesi[49,5,35], A. Mangilli[61], D. Marinucci[40], M. Maris[48], D. J. Marshall[73], P. G. Martin[8], E. Martínez-González[67*], S. Masi[36], S. Matarrese[34], F. Matthai[78], P. Mazzotta[39], J. D. McEwen[26], P. R. Meinhold[32], A. Melchiorri[36,52], L. Mendes[43], A. Mennella[37,50], M. Migliaccio[64,70], K. Mikkelsen[65], S. Mitra[55,68], M.-A. Miville-Deschênes[60,8], D. Molinari[49], A. Moneti[61], L. Montier[95,9], G. Morgante[49], D. Mortlock[56], A. Moss[88], D. Munshi[87], P. Naselsky[82,41], F. Nati[36], P. Natoli[35,4,49], C. B. Netterfield[21], H. U. Nørgaard-Nielsen[16], F. Noviello[69], D. Novikov[56], I. Novikov[82], S. Osborne[91], C. A. Oxborrow[16], F. Paci[86], L. Pagano[36,52], F. Pajot[60], D. Paoletti[49,51], F. Pasian[48], G. Patanchon[1], H. V. Peiris[26], O. Perdereau[71], L. Perotto[75], F. Perrotta[86], F. Piacentini[36], M. Piat[1], E. Pierpaoli[25], D. Pietrobon[68], S. Plaszczynski[71], E. Pointecouteau[95,9], D. Pogosyan[29], G. Polenta[4,47], N. Ponthieu[60,53], L. Popa[62], T. Poutanen[45,27,2], G. W. Pratt[73], G. Prézeau[10,68], S. Prunet[61,94], J.-L. Puget[60], J. P. Rachen[22,78], C. Räth[79], R. Rebolo[66,14,42], M. Reinecke[78], M. Remazeilles[60,1], C. Renault[75], A. Renzi[86], S. Ricciardi[49], T. Riller[78], I. Ristorcelli[95,9], G. Rocha[68,10], C. Rosset[1], A. Rotti[55], G. Roudier[1,72,68], J. A. Rubiño-Martín[66,42], B. Rusholme[57], M. Sandri[49], D. Santos[75], G. Savini[84], D. Scott[24], M. D. Seiffert[68,10], E. P. S. Shellard[11], T. Souradeep[55], L. D. Spencer[87], J.-L. Starck[73], V. Stolyarov[6,70,90], R. Stompor[1], R. Sudiwala[87], F. Sureau[73], P. Sutter[61], D. Sutton[64,70], A.-S. Suur-Uski[27,45], J.-F. Sygnet[61], J. A. Tauber[44], D. Tavagnacco[48,38], L. Terenzi[49], L. Toffolatti[19,67], M. Tomasi[50], M. Tristram[71], M. Tucci[17,71], J. Tuovinen[81], M. Türler[54], L. Valenziano[49], J. Valiviita[45,27,65], B. Van Tent[76], J. Varis[81], P. Vielva[67], F. Villa[49], N. Vittorio[39], L. A. Wade[68], B. D. Wandelt[61,94,33], I. K. Wehus[68], M. White[30], A. Wilkinson[69], D. Yvon[15], A. Zacchei[48], and A. Zonca[32]

(Affiliations can be found after the references)

Received XX, 2013; accepted XX, 2013

ABSTRACT

The two fundamental assumptions of the standard cosmological model — that the initial fluctuations are statistically isotropic and Gaussian — are rigorously tested using maps of the CMB anisotropy from the *Planck* satellite. The detailed results are based on studies of four independent estimates of the CMB that are compared to simulations using a fiducial ΛCDM model and incorporating essential aspects of the *Planck* measurement process. Deviations from isotropy have been found and demonstrated to be robust against component separation algorithm, mask and frequency dependence. Many of these anomalies were previously observed in the *WMAP* data, and are now confirmed at similar levels of significance (around 3σ). However, we find little evidence for non-Gaussianity with the exception of a few statistical signatures that seem to be associated with specific anomalies. In particular, we find that the quadrupole-octopole alignment is also connected to a low observed variance of the CMB signal. The dipolar power asymmetry is now found to persist to much smaller angular scales, and can be described in the low-ℓ regime by a phenomenological dipole modulation model. Finally, it is plausible that some of these features may be reflected in the angular power spectrum of the data which shows a deficit of power on the same scales. Indeed, when the power spectra of two hemispheres defined by a preferred direction are considered separately, one shows evidence for a deficit in power, whilst its opposite contains oscillations between odd and even modes that may be related to the parity violation and phase correlations also detected in the data. Whilst these analyses represent a step forward in building an understanding of the anomalies, a satisfactory explanation based on physically motivated models is still lacking.

Key words. Cosmology: observations, cosmic microwave background, isotropy, Gaussianity

1. Introduction

This paper, one of a set associated with the 2013 release of data from the *Planck*[1] mission (Planck Collaboration I 2013)

*Corresponding author: E. Martínez-González, email: martinez@ifca.unican.es

[1] *Planck* (`http://www.esa.int/Planck`) is a project of the European Space Agency (ESA) with instruments provided by two sci-

describes a set of studies undertaken to determine the statistical properties of the cosmic microwave background (CMB).

The standard cosmological model is well described by the *Friedmann-Lemaitre-Robertson-Walker* (FLRW) solution of the Einstein field equations. This model is characterized by a homogeneous and isotropic metric and an expanding scale factor of the Universe. At very early times it is hypothesized that the universe went through a period of accelerated expansion, the so-called *cosmological inflation*, driven by a hypothetical scalar field, the *inflaton*. During inflation the universe behaves as a de Sitter space, providing the conditions in which some of the present properties of the universe can be realized and specifically relaxing the problem of initial conditions. In particular, the seeds that gave rise to the present large-scale matter distribution via gravitational instability originated as quantum fluctuations of the inflaton about its vacuum state. These fluctuations in the inflaton produce energy perturbations which are distributed as a homogeneous and isotropic Gaussian random field. Linear theory relates those energy fluctuations to the CMB anisotropies, implying a distribution for the anisotropies very close to that of an isotropic Gaussian random field.

The scope of this work is to use *Planck* data to test the Gaussianity and near isotropy of the CMB in intensity as expected in the standard cosmology paradigm. Testing these fundamental properties is crucial for the validation of the standard cosmological scenario, and has profound implications for our understanding of the physical nature of the Universe and the initial conditions of structure formation. Moreover, the confirmation of the isotropic and Gaussian nature of the CMB is essential to justify the corresponding assumptions usually made in the estimation of the CMB power spectra, and other quantities to be obtained from the *Planck* data. Conversely, the detection of significant deviations from these assumptions that are not consistent with known systematic effects or foreground residuals would necessitate major revision of current methodological approaches for the derivation of the mission's many science results.

Significant deviations from Gaussianity are expected from non-linear processes that lead to secondary anisotropies, e.g. the integrated Sachs-Wolfe (ISW) effect and lensing. Indeed, these effects are the subject of two companion *Planck* papers (Planck Collaboration XVII 2013; Planck Collaboration XIX 2013, respectively). However, remarkably, a number of anomalies, by which we mean features of the observed sky that are not statistically consistent with the best-fit ΛCDM model, have been found in the *WMAP* data. Indeed, the *WMAP* team (Spergel et al. 2003) themselves initially proposed some intriguing discrepancies in the form of a lack of power on large angular scales. Further examples include an alignment of the low order multipoles some of which also indicate anomalously low amplitudes (Tegmark et al. 2003; Bielewicz et al. 2005; Land & Magueijo 2005a), a North-South asymmetry in both power spectra (Eriksen et al. 2004a; Hansen et al. 2009) and various measures of non-Gaussianity (Eriksen et al. 2004c, 2005; Räth et al. 2007a), parity asymmetry in the power spectrum corresponding to large angular scales (Kim & Naselsky 2010a) and a region of significant temperature decrement — the so-called *Cold Spot* (Vielva et al. 2004; Cruz et al. 2005).

Whilst *WMAP* have presented refutations of these anomalies, either by criticism of the robustness of the statistical methods employed (Bennett et al. 2011) or by associating them with systematic artefacts of the data processing that have been corrected in the nine-year data release (Bennett et al. 2012), *Planck* represents a unique opportunity to independently assess their existence. Its higher angular resolution and sensitivity and wider frequency range will allow a better understanding and removal of the Galactic and extragalactic foregrounds thus allowing a larger fraction of the sky to be useful for performing isotropy and Gaussianity analysis and to confirm and interpret those anomalies.

Throughout this paper, we quantify the significance of the test statistic in terms of the p-value. This is the probability of obtaining a test statistic at least as extreme as the observed one, under the assumption that the null hypothesis (i.e., Gaussianity and isotropy of the CMB) is true. In some tests, where it is well motivated to use only a one-tailed probability, the p-value is replaced by the corresponding upper or lower-tail probability. A low p-value is indicative of a tension between the data and the assumed statistical model (the null hypothesis). This can arise either when the assumed cosmological model is incorrect, if unknown or unmodelled aspects of the foreground emission or the measurement process exist, or as a result of a natural statistical fluctuation. The most interesting possibility, of course, is that a low p-value is an indication of new physics.

From the theoretical point of view, there are many variants of inflation that predict high levels of non-Gaussianity and new scenarios motivated by string and M-theory. In addition, there are many physical effects that might give rise to a deviation from isotropy or the presence of non-Gaussianity. Those deviations may be classified according to their physical nature and origin as follows: non-standard inflationary models, geometry and topology of the Universe, and topological defects. The main results from these areas, as well as the detailed descriptions of methodologies and of specific theoretically-motivated model constraints, are provided in the companion papers Planck Collaboration XXIV (2013), Planck Collaboration XXVI (2013), and Planck Collaboration XXV (2013).

This paper covers all relevant aspects related to the phenomenological study of the statistical isotropy and Gaussian nature of the CMB measured by the *Planck* satellite. It is organized as follows. Section 2 describes the *Planck* data used for the analyses. Section 3 explains the main characteristics of the simulations that constitute our reference set of Gaussian sky maps representative of the null hypothesis. In Sect. 4 the null hypothesis is tested with a number of standard tests that probe different aspects of non-Gaussianity. The *WMAP* anomalies are revisited in the light of the *Planck* data in Sect. 5. In Sect. 6 the implications of the found deviations of the null hypothesis on C_ℓ and cosmological parameters estimations are discussed. Finally, Sect. 7 provides the main conclusions of the paper.

2. Data description

In this paper, we utilise data from the *Planck*-2013 data release corresponding to the nominal period of the *Planck* mission. In part, this comprises sky maps at nine frequencies, with corresponding 'half-ring' maps that are generated by separating the data for a given pointing period into two halves, plus maps generated from data within the first and second survey periods. This set of maps allow a variety of consistency checks to be made, together with estimates of the instrumental noise contributions to analyses and limits on time-varying systematic artefacts. Full details are provided in papers Planck Collaboration II (2013); Planck Collaboration VI (2013).

entific consortia funded by ESA member states (in particular the lead countries France and Italy), with contributions from NASA, (USA) and telescope reflectors provided by a collaboration between ESA and a scientific conosrtium led and funded by Denmark

Our main results are based on the CMB maps resulting from sophisticated component separation algorithms applied to the frequency maps, as detailed in Planck Collaboration XII (2013). The four methods — `Commander-Ruler`, `NILC`, `SEVEM` and `SMICA` — are used to generate estimates of the CMB sky with an effective angular resolution of around 7′ or better, with accompanying symmetrised beam profiles, analysis masks, half-ring and survey maps. In general, the analyses presented here make use of a standardised common mask that merges those associated with the individual methods (this mask is listed in Table 1 as U73). This is a conservative approach and therefore, where appropriate, we manipulate the masks for use at lower resolution. Low resolution maps are required in some analyses and have been produced as follows. For resolutions $N_{\rm side}$=128–1024 the full resolution maps have been degraded using the `ud_degrade HEALPix` (Górski et al. 2005) routine. For degrading to even lower resolutions, $N_{\rm side}$ =16–64, a different procedure has been followed. Before degrading the maps to the final resolution using the ud_degrade routine as in the previous case, the full-resolution map is smoothed with a Gaussian kernel with a FWHM equal to three times the pixel size of the low resolution map that we want to produce.

In Table 1 we list the different masks that have been used in the analyses described in this paper. These masks have been produced at full resolution ($N_{\rm side}$=2048) and are described in papers Planck Collaboration XII (2013); Planck Collaboration XV (2013). The mask U73 is the most often used in this paper. However, for several applications the masks have been degraded to lower resolutions ($N_{\rm side}$ =1024, 512, 256, 128, 64, 32 and 16). The masks with resolutions $N_{\rm side}$=128–1024 have been degraded using the following procedure: first, the masks are degraded to their final resolution using the `ud_degrade HEALPix` routine, and then, a conservative approach is followed setting to zero any pixels with a value lower than 0.8. If the masks have to be degraded to even lower resolutions, $N_{\rm side}$ =16–64, the procedure that has been used is different. First, the full-resolution mask is smoothed with a Gaussian kernel with a FWHM equal to three times the pixel size of the low resolution mask that we want to produce. Then the mask is degraded using `ud_degrade` to their final resolution. Finally, those pixels with a value lower or equal than 0.5 have been set to zero and the rest have been set to 1. This criterion, less conservative than the one used for the higher resolution masks, is a compromise between minimizing the amount of sky that is being masked and the level of contamination left unmasked (we remark that in some cases the more conservative criterion of a 0.8 threshold has been also used for the lower resolutions, as stated in the corresponding analyses).

Table 1. List of the masks that have been used for the analyses described in this paper. All of them have been generated at $N_{\rm side}$ = 2048, and when needed, they have been degraded to a lower resolution as explained in the text. The CL masks have been constructed following the procedure described in Planck Collaboration XV (2013) but for different sky coverages.

Mask name	Sky coverage [% of unmasked pixels]
CS-SMICA89[1] . .	89.0
U73[1]	73.0
CG90[1]	90.0
CG80[1]	80.0
CG70[1]	70.0
CG60[1]	60.0
CL65[2]	65.1
CL58[2]	57.8
CL48[2]	48.0
CL37[2]	37.3
CL25[2]	24.7

[1]Planck Collaboration XII (2013)
[2] Planck Collaboration XV (2013)

3. Simulations

The derivation of results to be presented in this paper requires extensive simulations, essential aspects of which include: 1) modeling the *Planck* instrumental effects that affect the quality of the data, including instrumental noise and identified systematic effects, 2) replicating the foreground removal approach and estimating the extent of foreground residuals, and 3) modeling the intrinsic statistical properties, Gaussian or otherwise, of the CMB signals expected from specific models of the Universe.

The full focal plane (FFP6) simulations described in Planck Collaboration ES (2013) provide a complete realisation of the *Planck* mission capturing all characteristics of the scanning strategy, telescope, detector responses, and data reduction pipeline over the nominal mission period of 15.5 months. The Planck Sky Model (PSM) is used as input, encompassing the best current estimate of the microwave sky at *Planck* wavelengths including Galactic and extragalactic astrophysical foreground emission. The outputs include a complete set of maps for all detectors with accompanying half-ring and survey splits generated for a reference CMB sky. These have been used to test and validate various analysis tools, employed in turn to evaluate the CMB component separation algorithms as applied to the data set. This also allows an FFP6-based estimate of the foreground residuals remaining in the CMB sky after component separation to be evaluated, and their impact on various statistical estimators quantified.

An accompanying set of Monte Carlo simulations provides us with the reference set of Gaussian sky maps used for the null tests we employ. These simulations include `FEBeCoP` (Mitra et al. 2011) beam convolution at each of the *Planck* frequencies, which are then propagated through the various component separation pipelines using the same weights as derived from the *Planck* nominal mission data analysis. A fiducial CMB power spectrum has been adopted based on an analysis of the *Planck* data at an advanced, but not final stage of processing. Only small changes relative to the final *Planck* power spectrum presented in Planck Collaboration XV (2013); Planck Collaboration XVI (2013) are observed.

4. Are the primordial fluctuations Gaussian?

As has been previously established, it is of major interest to determine whether the statistical properties of the primordial CMB anisotropies correspond to an isotropic Gaussian random field. Recent attempts to test this hypothesis have mainly relied on the *WMAP* data that have less sensitivity and cover a narrower frequency interval. *Planck* represents a unique opportunity to probe fundamental statistical properties of the Universe with cosmic variance limited sensitivity up to $\ell \approx 2000$ and minimum foreground contamination.

There is no unique signature of non-Gaussianity, however, different tests can allow us to probe different types of non-Gaussianity. As a consequence, it is important to subject the data to a variety of tests, and we do so in this section using a number of non-parametric tools. Specific signatures of non-Gaussianity are sought in three companion papers — Planck

Collaboration XXIV (2013); Planck Collaboration XXV (2013); Planck Collaboration XXVI (2013).

Any isotropic and continuous random field, $T(\boldsymbol{x})$ on the sphere can be written in terms of the following spectral representation:

$$T(\boldsymbol{x}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\boldsymbol{x}), \tag{1}$$

where $\boldsymbol{x}$ is a unit direction vector, $Y_{\ell m}(\cdot)$ the spherical harmonics and

$$a_{\ell m} = \int \mathrm{d}\boldsymbol{x} T(\boldsymbol{x}) Y^*_{\ell m}(\boldsymbol{x}), \tag{2}$$

$m = 0, \pm 1, ..., \pm\ell, \quad \ell = 1, 2, ...$

For a Gaussian field with uncorrelated phases, each $a_{\ell m}$ coefficient will be independent with a zero mean Gaussian distribution:

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = \delta^{\ell}_{\ell'} \delta^{m}_{m'} C_\ell \tag{3}$$

where δ is the Kronecker delta and C_ℓ is the angular power spectrum. Note that for a Gaussian and isotropic random field, the angular power spectrum provides a complete characterization of its statistical distribution.

In this paper, we examine the goodness-of-fit of the data to the *Planck* best-fit fiducial CMB model, which constitutes our null hypothesis. The methods adopted constitute a broad range of statistical tools that allow the study of complementary statistical properties of the null hypothesis in both the real and harmonic space data representations. Claims of either consistency with the fiducial *Planck* cosmological model or of evidence for non-Gaussianity must be demonstrably robust to data selection and specifics of the data analysis. Residuals from the diffuse Galactic foreground are likely to be non-Gaussian in nature, and point-sources can be a source of non-Gaussianity on small angular scales. In addition, the analysis of multifrequency data must be considered in order to confirm that any claimed non-Gaussianity has a thermal (cosmological) spectrum. Moreover, the combined ISW-lensing effect produces secondary anisotropies that significantly deviate from Gaussianity and whose effect has been detected in the *Planck* data (Planck Collaboration XIX 2013). This non-Gaussian effect has to be considered when testing the null hypothesis.

We address these issues by analysing the cosmologically interesting subset of *Planck* frequency channels. Specifically, we analyse the uncorrected sky maps at 70, 100, 143 and 217 GHz as a function of Galactic mask to assess the likely contamination due to Galactic foregrounds. These tests have direct relevance for the *Planck* likelihood approach described in Planck Collaboration XV (2013) and the parameter estimation results presented in Planck Collaboration XVI (2013). We then consider the foreground cleaned versions of these maps generated by the SEVEM algorithm (see Planck Collaboration XII 2013). Such a comparison also allows a semi-independent cross-check of the cosmological signal seen by *Planck* LFI (70 GHz) and HFI (100, 143, 217 GHz). Although the cosmological content of the cleaned LFI and HFI data sets are independent, the cleaning makes use of difference maps generated from the remaining *Planck* frequency bands. Nevertheless, since the calibration and beam responses of the data are well understood over the ful range of frequencies, there will be no leakage of cosmological signal between the instrument specific frequencies.

We then continue with analyses of the CMB sky estimates provided by four component separation approaches (Commander-Ruler, NILC, SEVEM, and SMICA) described in Planck Collaboration XII (2013), together with the corresponding mask appropriate for these methods. The largest sky area possible should be used for definitive statements about Gaussianity since, in the absence of foreground residuals or systematic artefacts, it represents a superior sample of the Universe. Conversely, overly conservative sky cuts suffer from a loss of information.

Table 2. Lower tail probablity for the variance, skewness and kurtosis estimators at $N_{\rm side} = 2048$, using the U73 mask and four different component separation methods.

Method	Variance	Skewness	Kurtosis
C-R	0.021	0.189	0.416
NILC	0.020	0.191	0.392
SEVEM	0.014	0.206	0.419
SMICA	0.017	0.189	0.419

4.1. One dimensional moments

In this section we perform some of the simplest Gaussianity tests, such as comparing the sample skewness and kurtosis of the data with simulations. The skewness, γ, and kurtosis, κ, of a random variable, X, are defined as follows:

$$\gamma(X) = \frac{\langle X - \langle X \rangle \rangle^3}{(\mathrm{Var}(X))^{3/2}} \tag{4}$$

$$\kappa(X) = \frac{\langle X - \langle X \rangle \rangle^4}{(\mathrm{Var}(X))^2} - 3 \tag{5}$$

The skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable. Qualitatively, a positive (negative) skew indicates that the tail on the right (left) side of the probability density function is longer than the left (right) side. A zero value indicates that the values are relatively evenly distributed on both sides of the mean, typically but not necessarily implying a symmetric distribution. The kurtosis is a measure of the peakedness of the distribution and the heaviness of its tail. A distribution with positive (negative) excess kurtosis indicates that the distribution has a more acute (wider) peak around the mean and fatter (thinner) tails. Normal random variables have zero skewness and kurtosis.

The sample variance is also considered in this section as a further consistency test, although it is not a normality test statistic.

We begin by analysing the full resolution combined maps, applying the U73 mask for the four different component separation methods. The results for the variance, skewness and kurtosis estimators are shown in Table 2. All four methods show similar results. The data are consistent with simulations for the skewness and kurtosis estimators, whereas the variance is anomalously low. This inconsistency was already reported for the *WMAP* data in Monteserín et al. (2008) and Cruz et al. (2011) at resolution $N_{\rm side} = 256$ for a mask allowing slightly less sky coverage.

The mask dependence of our results is studied by recalculating the estimators using the CL58 and CL37 masks which allow sky fractions of $f_{\rm sky} = 58\%$ and $f_{\rm sky} = 37\%$ respectively. The SMICA cleaned maps at full resolution are considered. The most significant lower tail probability is obtained for the CL58 mask as can be seen in Table 3. The lower tail probabilities show a small dependence on the mask used, which could indicate either the presence of Galactic foreground residuals with larger

Table 3. Lower tail probablity for the variance, skewness and kurtosis estimators at $N_{\rm side} = 2048$, for the SMICA method, using different masks.

Mask	Variance	Skewness	Kurtosis
U73, $f_{\rm sky}$ =73%	0.017	0.189	0.419
CL58, $f_{\rm sky}$ =58%	0.003	0.170	0.363
CL37, $f_{\rm sky}$ =37%	0.030	0.314	0.266
Ecliptic North, $f_{\rm sky}$ =36%	0.001	0.553	0.413
Ecliptic South, $f_{\rm sky}$ =37%	0.483	0.077	0.556
Galactic North, $f_{\rm sky}$ =37%	0.001	0.788	0.177
Galactic South, $f_{\rm sky}$ =36%	0.592	0.145	0.428

Table 4. Lower tail probablity for the variance, skewness and kurtosis estimators at $N_{\rm side} = 2048$, for the SEVEM cleaned maps at different frequencies.

Map	Variance	Skewness	Kurtosis
100 GHz	0.023	0.195	0.488
143 GHz	0.014	0.221	0.460
217 GHz	0.025	0.196	0.481

sky coverage, or the increase of the sampling variance, and consequently a less significant probability, when a smaller fraction of the sky is considered.

In order to identify any foreground contamination, the frequency dependence of our estimators is analysed. We use the SEVEM cleaned maps and the U73 mask. Note that as the 70 GHz full resolution noise is high we do not consider 70 GHz in this comparison. As the 100 GHz noise is not negligible we estimate the variance taking into account the noise dispersion as described in Cruz et al. (2011). The results are similar to those found for the combined map, as can be seen in Table 4. There is a small frequency dependence since the 100 GHz and 217 GHz maps show slightly higher variance and kurtosis than the 143 GHz map. However the 143 GHz map has a dominant contribution to the combined map, hence the foreground residuals in the combined map are likely to be small. The lower tail probabilities for the variance at 100 GHz, 143 GHz, and 217 GHZ are respectively 0.021, 0.014, 0.025, whereas the skewness and kurtosis are compatible with simulations.

We also reanalyse the SMICA data and simulations considering independently the northern and southern ecliptic hemispheres outside the U73 mask. A clear asymmetry is found in the variance, with an anomalously low value found in the northern hemisphere, as seen in Table 3.

The results for different resolutions using the U73 mask are shown in Table 5. Note that the $N_{\rm side} = 2048$ and 512 U73 masks have $f_{\rm sky} = 73\%$, while the low resolution masks at $N_{\rm side} = 64$, 32, and 16 have $f_{\rm sky} = 78\%$. The variance is anomalously low at all the considered resolutions, whereas at low resolutions, the skewness is anomalously low and the kurtosis anomalously high. These results will be further analysed in Sect. 5.2. However, it is clear that, except on the largest angular scales, there is no evidence for non-Gaussian behaviour in the data using these simple statistical measures.

Table 5. Lower tail probablity for the variance, skewness and kurtosis estimators at different resolutions, for the four component separation methods, using the U73 mask.

Method	Variance	Skewness	Kurtosis
	$N_{\rm side} = 2048$		
C-R	0.021	0.189	0.416
NILC	0.020	0.191	0.392
SEVEM	0.014	0.206	0.419
SMICA	0.017	0.189	0.419
	$N_{\rm side} = 512$		
C-R	0.017	0.207	0.368
NILC	0.017	0.198	0.390
SEVEM	0.013	0.218	0.408
SMICA	0.014	0.196	0.390
	$N_{\rm side} = 64$		
C-R	0.011	0.041	0.935
NILC	0.011	0.041	0.935
SEVEM	0.008	0.058	0.900
SMICA	0.011	0.041	0.943
	$N_{\rm side} = 32$		
C-R	0.020	0.015	0.968
NILC	0.019	0.016	0.960
SEVEM	0.012	0.026	0.932
SMICA	0.019	0.016	0.967
	$N_{\rm side} = 16$		
C-R	0.023	0.013	0.974
NILC	0.022	0.022	0.972
SEVEM	0.019	0.022	0.964
SMICA	0.027	0.021	0.982

4.2. *N*-pdf analysis

Under the assumption of Gaussianity, the *N*-probability density function (*N*-pdf) is given by a multivariate Gaussian function:

$$f(\boldsymbol{T}) = \frac{1}{(2\pi)^{N_{pix}/2}\det \mathsf{C}^{1/2}} \exp -\frac{1}{2}\left(\boldsymbol{T}\mathsf{C}^{-1}\boldsymbol{T}^{\rm T}\right), \quad (6)$$

where $\boldsymbol{T}$ is a vector formed from the measured temperatures $T(\boldsymbol{x})$ over all positions allowed by the applied mask, N_{pix} is the number of pixels in the vector, C is the covariance of the Gaussian field (of size $N_{\rm pix} \times N_{\rm pix}$).

Unfortunately, the calculation of $\boldsymbol{T}\mathsf{C}^{-1}\boldsymbol{T}^{\rm T}$ is computationally unfeasible for the full *Planck* resolution at HEALPix $N_{\rm side} = 2048$. At a lower resolution, the problem is tractable, and the noise level can also be considered negligible compared to the CMB signal. That implies that under the assumption of isotropy the covariance matrix C is fully defined by the *Planck* angular power spectrum (C_ℓ):

$$\mathsf{C}_{ij} = \sum_{\ell=0}^{\ell_{\rm max}} \frac{2\ell+1}{4\pi} C_\ell b_\ell^2 P_\ell\left(\cos\theta_{ij}\right), \quad (7)$$

where C_{ij} is the covariance between pixels i and j, and θ_{ij} is angle between them, P_ℓ are the Legendre polynomials, b_ℓ is an effective window function associated with the Nside resolution, and $\ell_{\rm max}$ is the maximum multipole probed.

Under the multivariate Gaussian hypothesis, the argument on the exponential in equation 6 should follow a χ^2 distribution with $N_{\rm pix}$ degrees of freedom, or, equivalently (for $N_{\rm pix} >> 1$) a normal distribution $\mathcal{N}\left(N_{\rm pix}, \sqrt{N_{\rm pix}}\right)$.

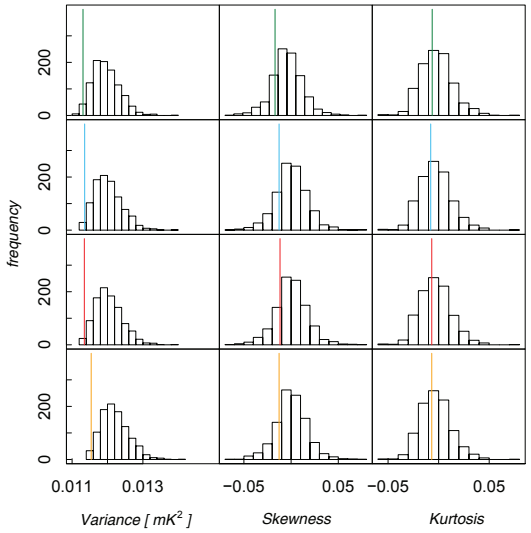


Fig. 1. Variance, skewness and kurtosis for the combined map of the four different component separation methods. From top row to bottom row `C-R`, `NILC`, `SEVEM`, `SMICA`.

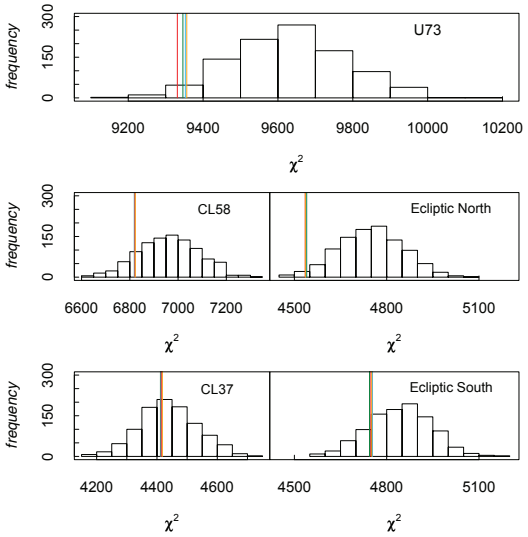


Fig. 2. N-pdf χ^2 for the U73 mask, CL58, CL37, ecliptic North and ecliptic South. The different colours represent the four component separation methods, namely `C-R` (green), `NILC` (blue), `SEVEM` (red), and `SMICA` (orange).

Table 6. Lower tail probablity for the N-pdf, using different masks.

Mask	C-R	NILC	SEVEM	SMICA
U73, $f_{\rm sky}$ =78%	0.027	0.028	0.019	0.030
CL58, $f_{\rm sky}$ =58%	0.137	0.137	0.147	0.146
CL37, $f_{\rm sky}$ =37%	0.409	0.415	0.420	0.436
Ecliptic North, $f_{\rm sky}$ =39%	0.024	0.022	0.021	0.021
Ecliptic South, $f_{\rm sky}$ =39%	0.170	0.196	0.183	0.193

Table 7. Frequency dependence of the lower tail probablity for the N-pdf, using different masks.

Mask	70 GHz	100 GHz	143 GHz	217 GHz
U73, $f_{\rm sky}$ =78%	0.037	0.058	0.013	0.124
CL58, $f_{\rm sky}$ =58%	0.169	0.123	0.111	0.169
CL37, $f_{\rm sky}$ =37%	0.422	0.366	0.376	0.386
Ecliptic North, $f_{\rm sky}$ =39%	0.028	0.050	0.015	0.083
Ecliptic South, $f_{\rm sky}$ =39%	0.225	0.233	0.166	0.330

We begin by analysing the χ^2 quantity for low resolution maps at $N_{\rm side} = 32$ and filtering with a 5° FWHM Gaussian. $1\,\mu K$ uncorrelated regularization noise is added to the covariance matrix before inverting it. Regularization noise realizations are added to the data and simulations for consistency (see Eriksen et al. 2007b, for more details).

We analyse the four cleaned data maps, applying the common, CL58 and CL37 masks. The admitted fraction of the sky is respectively 78%, 58% and 37%. The northern and southern ecliptic hemispheres outside the U73 mask are also considered. The results are shown in Fig. 2 and Table 6. In the U73 mask case, the lower tail probabilities are low. Applying the two CL58 and CL37 masks that permit less sky coverage, the data are consistent with simulations. The low χ^2 value appears to be localised in the northern ecliptic hemisphere. These results are directly comparable to the anomalous variance mentioned in Sect. 4.1. Note that the four maps show similar values, but the differences are larger when using the U73 mask. This could indicate the presence of some residual foreground contamination near the Galactic plane. Therefore, the frequency dependence of our estimator is analysed in order to identify any possible foreground contamination. The results are shown in Fig. 3 and Table 7. A moderate frequency dependence is found when using the U73 mask, which could indicate the presence of some foreground residuals near the Galactic plane. The frequency dependence of the results vanishes when using the CL58 and CL37 masks that exclude more of the sky from analysis.

4.3. N-point correlation functions

In this section we present tests of the non-Gaussianity of the *Planck* CMB maps using real-space N-point correlation functions. While harmonic-space methods are often preferred over real-space methods for studying primordial fluctuations, real-space methods may have an advantage with respect to systematics and foregrounds, since such effects are usually localized in real space. It is therefore important to analyse the data in both spaces in order to highlight different features.

An N-point correlation function is by definition the average product of N temperatures, measured in a fixed relative orientation on the sky,

$$C_N(\theta_1, \ldots, \theta_{2N-3}) = \left\langle \Delta T(\hat{\mathbf{n}}_1) \cdots \Delta T(\hat{\mathbf{n}}_N) \right\rangle, \quad (8)$$

where the unit vectors $\hat{\mathbf{n}}_1, \ldots, \hat{\mathbf{n}}_N$ span an N-point polygon on the sky. By assuming statistical isotropy, the N-point functions are only functions of the shape and size of the N-point poly-

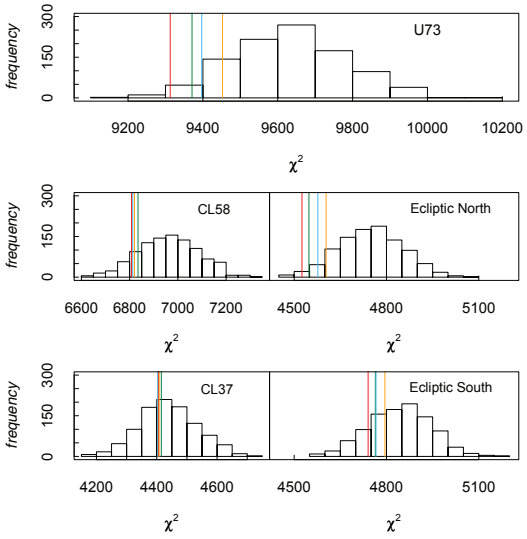


Fig. 3. Frequency dependence for 70 GHz (green), 100 GHz (blue), 143 GHz (red) and 217 GHz (orange), and different masks.

gon, and not on its particular position or orientation on the sky. Hence, the smallest number of parameters that uniquely determines the shape and size of the N-point polygon is $2N - 3$. In practice, the functions are estimated by simple product averages over all sets of N pixels fulfilling the geometric requirements set by $\theta_1, \ldots, \theta_{2N-3}$ characterising the shape and size of the polygon

$$C_N(\theta_1, \ldots, \theta_{2N-3}) = \frac{\sum_i \left(w_1^i \cdots w_N^i\right)\left(\Delta T_1^i \cdots \Delta T_N^i\right)}{\sum_i w_1^i \cdots w_N^i} \,. \quad (9)$$

Pixel weights $w_1^i, \cdots, w_N^i$ can be introduced in order to reduce noise or mask boundary effects. Here they represent masking by being set to 1 for included pixels and to 0 for excluded pixels.

The main difficulty with computing N-point functions is their computational scaling. The number of independent pixel combinations scales as $O(N_{\rm pix}^N)$, and for each combination of N pixels, $2N - 3$ angular distances must be computed to uniquely determine the properties of the corresponding polygon. Computing the full N-point function for $N > 2$ and $N_{\rm pix} \gtrsim 10^5$ is therefore computationally challenging. However, it is not necessary to include all possible N-point configurations in order to produce interesting results. For instance, one may focus only on small angular scales, or on configurations with some special symmetry properties. By using the methods described by Eriksen et al. (2004b), the computational expense then becomes tractable, since no CPU time is spent on excluded configurations. In this paper several such subsets are computed, covering three distinct ranges of scales, namely small (up to 3°), intermediate (up to 10°) and large angular scales (the full range between 0° and 180°). The shapes of considered polygons selected for the analysis are the pseudo-collapsed and equilateral configurations for the 3-point function, and the the rhombic configuration composed of two equilateral triangles for the 4-point function. In the following, all results refer to the reduced 4-point function, i.e., corrected for the Gaussian contribution due to the Wick's theorem. The size of the polygons is parametrised by the length of the longer side of the triangle in the case of the pseudo-collapsed configuration, and the length of the side for the equilateral triangle and rhombus.

We analyse the CMB estimates downgraded to $N_{\rm side} = 64$ and $N_{\rm side} = 512$ as well as at the original resolution of $N_{\rm side} = 2048$. In the case of the analysis at $N_{\rm side} = 64$ the maps were additionally smoothed with FWHM of 165′(three times the pixel size for the downgraded map). Due to computational limitations, an analysis is possible on the full sky only in the case of resolution $N_{\rm side} = 64$. For the higher resolution maps, we perform the analysis on a set of non-overlapping discs. For $N_{\rm side} = 512$ we uniformly retain, after masking, part of the sky with approximately 100 discs of radius 10°. Analogously to the analysis by Eriksen et al. (2005), we consider two disc sets A and B with a relative offset between their grids such that the centres of the discs of set B are located in parts of the sky not covered by disc set A (see Fig. 4). For studies at the original resolution $N_{\rm side} = 2048$, we restrict the analysis to 20 discs with a radius of 3° located randomly on an unmasked part of the sky (Fig. 4).

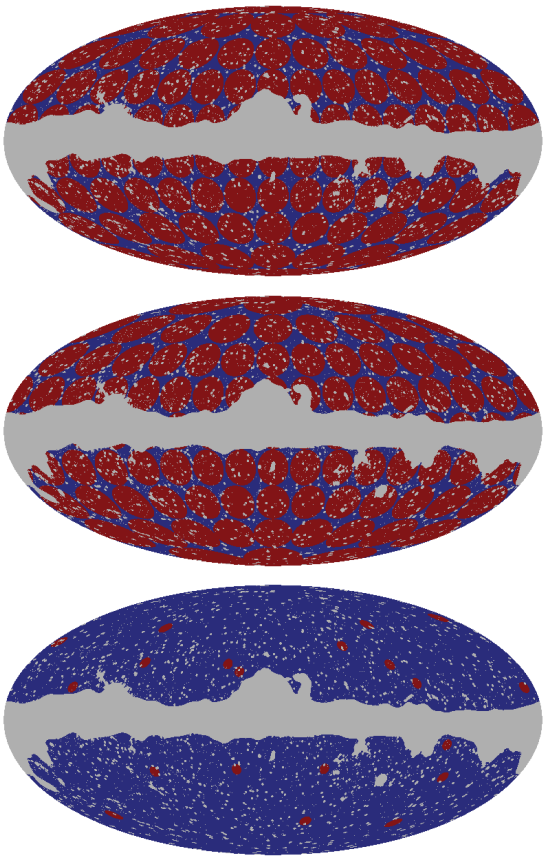


Fig. 4. Two sets of discs, A and B, each of radius 10° for the $N_{\rm side} = 512$ CMB estimates (upper and middle figure, respectively) and a set of 20 randomly placed discs of radius 3° superimposed on the U73 mask (blue region) for the CMB estimates at $N_{\rm side} = 2048$ (lower figure).

As in Eriksen et al. (2005), we consider the N-point correlation functions averaged over the disc sets. In order to minimize correlations between the discs, we subtract from the maps at resolutions $N_{\rm side} = 512$ and $N_{\rm side} = 2048$ the best-fit multipoles computed for the ranges $\ell \leq 18$ and $\ell \leq 60$, respectively. This procedure corresponds in practice to a high-pass filtering of the maps.

The low resolution versions of the U73 mask described earlier were used as required. Residual monopole and dipole contributions were then removed from the maps.

A simple χ^2 test is chosen to quantify the degree of agreement between the simulations and the observations, where χ^2 as usual is defined by

$$\chi^2 = \sum_{i,j=1}^{N_{\rm bin}} \left(C_N(\theta_i) - \left\langle C_N(\theta_i)\right\rangle\right) M_{ij}^{-1} \left(C_N(\theta_j) - \left\langle C_N(\theta_j)\right\rangle\right) . \quad (10)$$

Here $C_N(\theta_i)$ is the N-point correlation function for $i-th$ bin of separation angle, θ_i, $\left\langle C_N(\theta_i)\right\rangle$ is the corresponding average from the Monte Carlo (MC) ensemble, and

$$M_{ij} = \frac{1}{N_{\rm sim}} \sum_{k=1}^{N_{\rm sim}} \left(C_N^{(k)}(\theta_i) - \left\langle C_N(\theta_i)\right\rangle\right)\left(C_N^{(k)}(\theta_j) - \left\langle C_N(\theta_j)\right\rangle\right) \quad (11)$$

is the covariance matrix. Although the inverse of the covariance matrix constructed from MC simulations can be biased, it is relatively small for 1000 simulations and has a negligible impact on the significance levels estimated from the simulations, as described below.

This statistic is optimized for studying Gaussian distributed data. However, usually it also works quite well for mildly non-Gaussian distributions, and in particular symmetric ones. Nevertheless, as for any statistic constructed from MC simulations, it can also be used for non-Gaussian and asymmetrically distributed data. Below, we quote the significance level in terms of the fraction of simulations with a larger χ^2 value than the observed map.

We analyse the mask dependence of the non-Gaussianity of the maps using the pseudo-collapsed 3-point correlation function. The function averaged over disc set A is shown in Fig. 5. The corresponding probabilities of obtaining values of the χ^2 statistic for the concordance ΛCDM model at least as large as the observed values are given in Table 8.

Table 8. Probabilities of obtaining values of the χ^2 statistic for the concordance ΛCDM model at least as large as the observed values of the statistic for the raw 143 GHz (first row) and foreground corrected 143 GHz SEVEM CMB maps (second row).

$f_{\rm sky}$	0.6	0.7	0.8	0.9
Raw	0.907	0.889	0.563	0.000
SEVEM	0.959	0.959	0.905	0.940

In summary, the pseudo-collapsed 3-point function does not show any significant deviation from Gaussianity for the raw 143 GHz map masked with the CG60 ($f_{\rm sky} = 0.6$) and CG70 ($f_{\rm sky} = 0.7$) masks. To a lesser extent, this is true also for the CG80 ($f_{\rm sky} = 0.8$) mask. We do not see any significant deviation for any of the analysed masks after cleaning the 143 GHz map using the SEVEM method.

The correlation functions for the four component separation methods and resolution parameters $N_{\rm side} = 64$, $N_{\rm side} = 512$ and $N_{\rm side} = 2048$ are shown in Fig. 6, Fig. 7 (disc set A), Fig. 8 (disc set B) and Fig. 9, respectively. The probabilities of obtaining values of the χ^2 statistic for the *Planck* fiducial ΛCDM model at least as large as the observed values are given in the Tables 9, 10 and 11, respectively.

The results show consistency between the CMB maps estimated using the different component separation methods. We did not find statistically significant deviations of the CMB maps

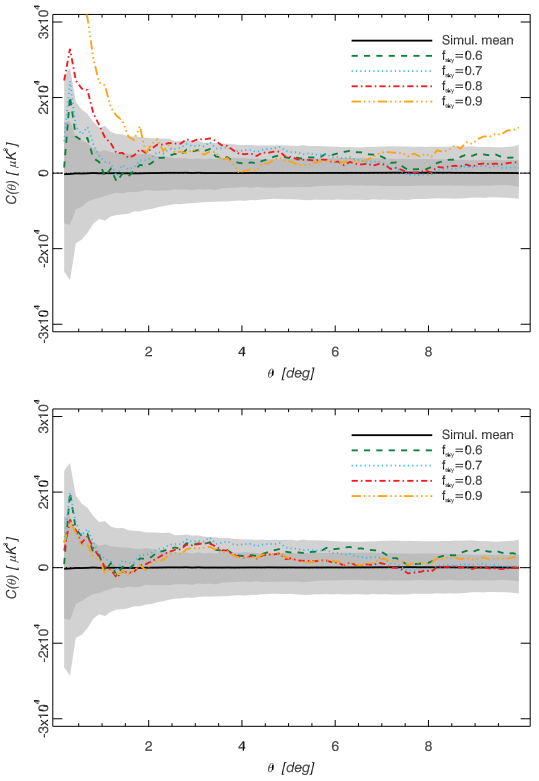


Fig. 5. The pseudo-collapsed 3-point function averaged over disc set A for the raw (upper figure) and SEVEM foreground corrected (lower figure) 143 GHz map at $N_{\rm side} = 512$. Estimates of the multipoles for $\ell \leq 18$ are removed from the sky maps. The black solid line indicates the mean for 1000 MC simulations and the shaded dark and light grey regions indicate the 68% and 95% confidence regions, respectively, for the CG90 ($f_{\rm sky} = 0.9$) mask.

Table 9. Probabilities of obtaining values for the χ^2 statistic of the N-point functions shown in Fig. 6 for the *Planck* fiducial ΛCDM model at least as large as the observed values of the statistic for the *Planck* CMB maps with resolution parameter $N_{\rm side} = 64$ estimated using the C-R, NILC, SEVEM and SMICA methods.

	C-R	NILC	SEVEM	SMICA
2-pt.	0.883	0.859	0.884	0.855
pseudo-coll. 3-pt.	0.922	0.918	0.945	0.908
equil. 3-pt.	0.962	0.966	0.978	0.968
4-pt.	0.975	0.977	0.979	0.977

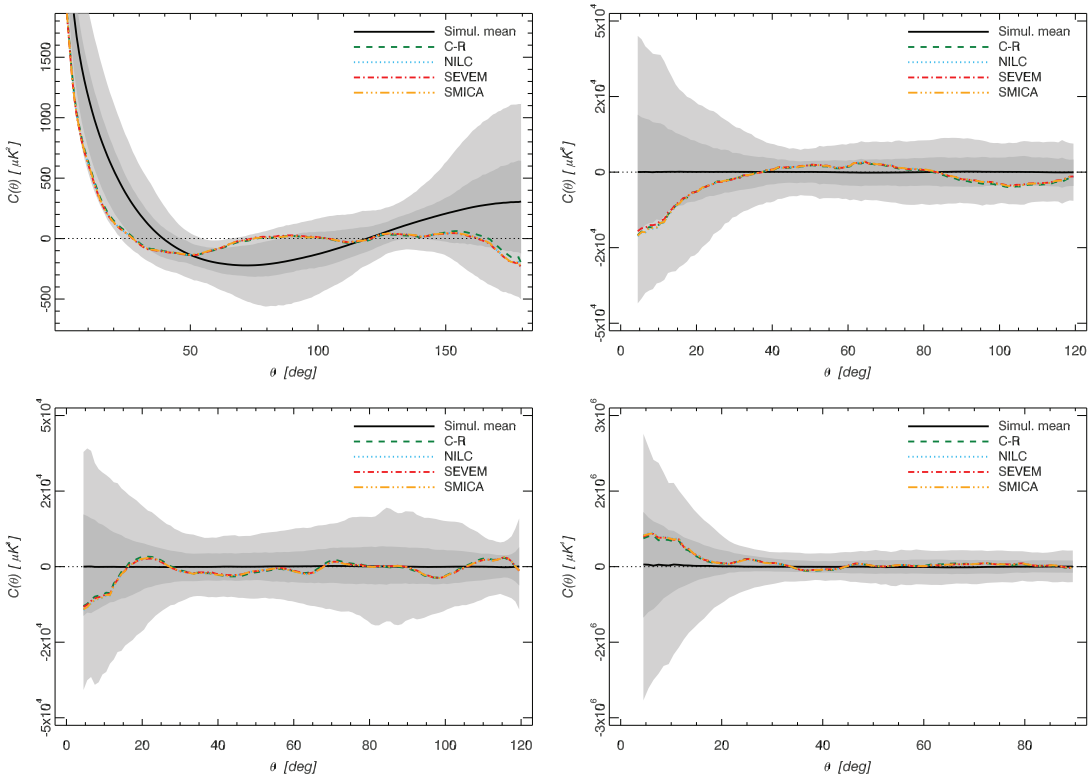


Fig. 6. The 2-point (upper left), pseudo-collapsed (upper right), equilateral (lower left) 3-point and reduced rhombic 4-point (lower right) functions for the $N_{\rm side} = 64$ CMB estimates. The black solid line indicates the mean for 1000 MC simulations and the shaded dark and light grey regions indicate the 68% and 95% confidence regions, respectively.

Table 10. Probabilities of obtaining values of the χ^2 statistic of the N-point functions shown in Figs. 7 and 8 for the *Planck* fiducial ΛCDM model at least as large as the observed values of the statistic for *Planck* CMB maps with resolution parameter $N_{\rm side} = 512$ estimated using the C-R, NILC, SEVEM and SMICA methods.

	C-R	NILC	SEVEM	SMICA
Two-point function				
A set	0.858	0.902	0.886	0.904
B set	0.351	0.370	0.404	0.376
Pseudo-collapsed three-point function				
A set	0.568	0.565	0.651	0.603
B set	0.483	0.526	0.550	0.540
Equilateral three-point function				
A set	0.004	0.032	0.045	0.043
B set	0.452	0.485	0.443	0.479
Rhombic four-point function				
A set	0.104	0.102	0.102	0.107
B set	0.521	0.569	0.537	0.579

Table 11. Probabilities of obtaining values of the χ^2 statistic of the N-point functions shown in Fig. 9 for the *Planck* fiducial ΛCDM model at least as large as the observed values of the statistic for *Planck* CMB maps with resolution parameter $N_{\rm side} = 2048$ estimated using the C-R, NILC, SEVEM, and SMICA methods.

	C-R	NILC	SEVEM	SMICA
2-pt.	0.335	0.474	0.573	0.497
pseudo-coll. 3-pt.	0.522	0.463	0.469	0.448
equil. 3-pt.	0.853	0.789	0.819	0.796
4-pt.	0.532	0.534	0.579	0.526

from Gaussianity for any of the analysed scales. However, it is clear that the CMB maps smoothed and downgraded to $N_{\rm side} = 64$ show the largest deviation, especially for the 4-point correlation function, in comparison to the intermediate and small angular scale analyses.

4.4. Minkowski functionals

Minkowski functionals (Minkowski 1903, hereafter MFs) describe the morphology of fields in any dimension and have long

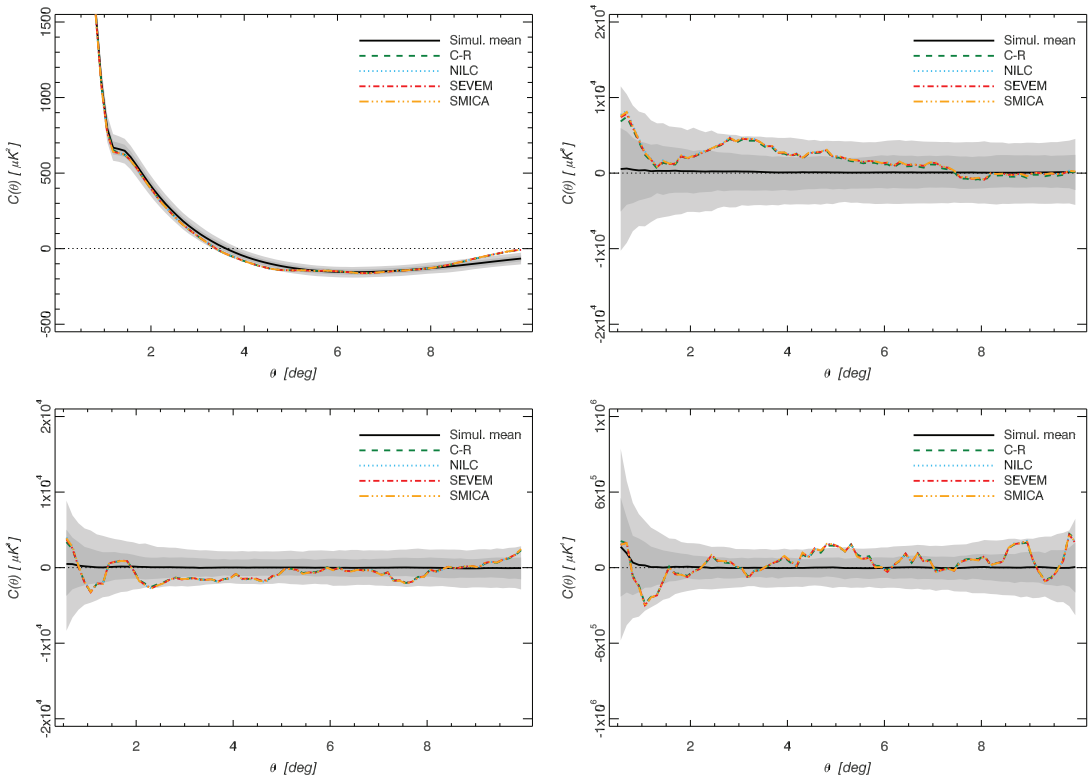


Fig. 7. The 2-point (upper left), pseudo-collapsed (upper right), equilateral (lower left) 3-point and reduced rhombic 4-point (lower right) functions averaged over disc set A for the $N_{\rm side} = 512$ CMB estimates. Estimates of the multipoles for $\ell \leq 18$ are removed from the sky maps. The black solid line indicates the mean for 1000 MC simulations and the shaded dark and light grey regions indicate the 68% and 95% confidence regions, respectively.

been used as estimators of non-Gaussianity and anisotropy in the CMB (see e.g., Gott et al. 1990; Mecke et al. 1994; Schmalzing & Gorski 1998; Komatsu et al. 2003; Eriksen et al. 2004c; Curto et al. 2007; De Troia et al. 2007; Spergel et al. 2007; Curto et al. 2008; Hikage et al. 2008; Komatsu et al. 2009). They are additive for disjoint regions of the sky and invariant under rotations and translations. Traditionally in the literature, the contours are defined by a threshold ν, usually given in units of the sky standard deviation (σ_0). We compute MFs for the regions colder and hotter than a given threshold ν. Thus, the three MFs, the area $V_0(\nu) = A(\nu)$, the perimeter $V_1(\nu) = C(\nu)$ and the genus $V_2(\nu) = G(\nu)$, are defined respectively as:

$$V_0(\nu) = A(\nu) = \frac{N_\nu}{N_{\rm pix}}, \tag{12}$$

$$V_1(\nu) = C(\nu) = \frac{1}{4A_{\rm tot}} \sum_i S_i, \tag{13}$$

$$V_2(\nu) = G(\nu) = \frac{1}{2\pi A_{\rm tot}} \Big(N_{\rm hot} - N_{\rm cold} \Big), \tag{14}$$

where N_ν is the number of pixels where $\Delta T/\sigma_0 > \nu$, $N_{\rm pix}$ is the total number of available pixels, $A_{\rm tot}$ is the total area of the available sky, $N_{\rm hot}$ is the number of compact hot spots, $N_{\rm cold}$ is the number of compact cold spots and S_i is the contour length of each hot spot. We construct a fourth functional $V_3(\nu) = N_{\rm cluster}(\nu)$ which corresponds to $N_{\rm cold}$ for negative ν and $N_{\rm hot}$ for positive ν (Ducout et al. 2012). Analytical expressions for a Gaussian random field can be derived in terms of ν (see e.g. Vanmarcke 1983; Matsubara 2010) and give the following,

$$V_k(\nu) = A_k v_k(\nu), \tag{15}$$

with

$$v_k(\nu) = \exp(-\nu^2/2) H_{k-1}(\nu), \quad k \leq 2 \tag{16}$$

$$v_3(\nu) = \frac{e^{-\nu^2}}{\mathrm{erfc}\left(\nu/\sqrt{2}\right)}, \tag{17}$$

and

$$H_n(\nu) = \mathrm{e}^{\nu^2/2} \left(-\frac{\mathrm{d}}{\mathrm{d}\nu} \right)^n \mathrm{e}^{-\nu^2/2}. \tag{18}$$

The amplitude A_k depends only on the shape of the power spectrum C_ℓ:

$$A_k = \frac{1}{(2\pi)^{(k+1)/2}} \frac{\omega_2}{\omega_{2-k}\omega_k} \left(\frac{\sigma_1}{\sqrt{2}\sigma_0} \right)^k, \quad k \leq 2 \tag{19}$$

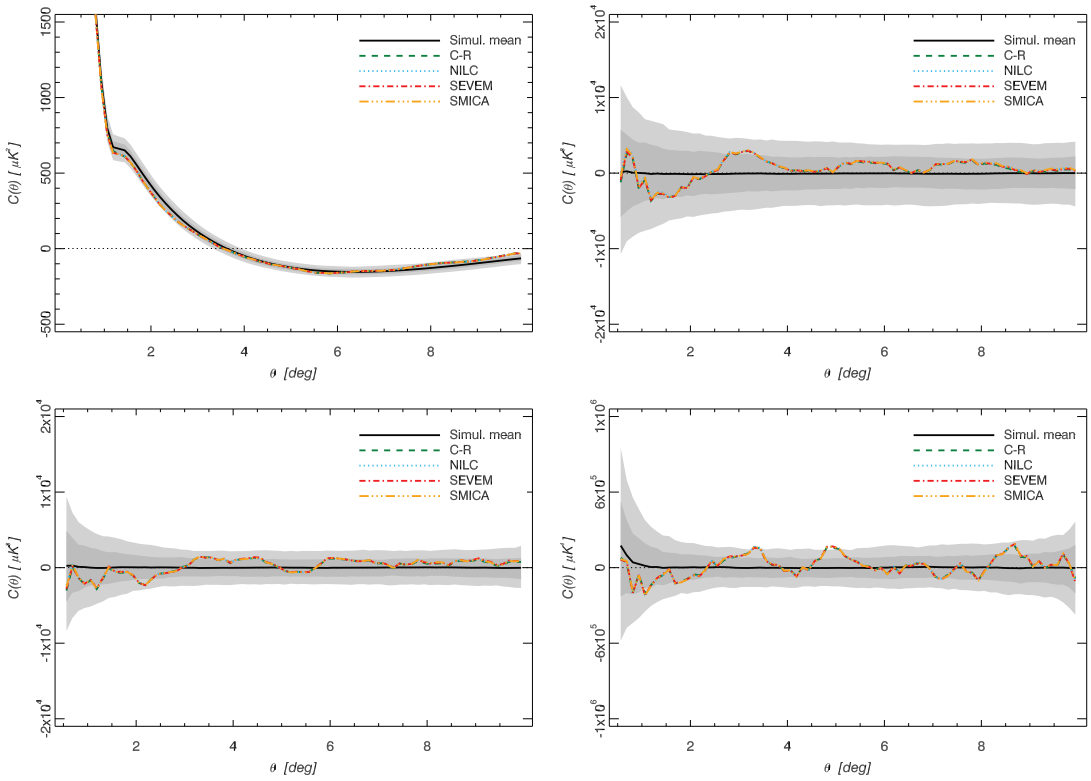


Fig. 8. As Fig. 7 for disc set B.

$$A_3 = \frac{2}{\pi}\left(\frac{\sigma_1}{\sqrt{2}\sigma_0}\right)^2 \quad (20)$$

where $\omega_k \equiv \pi^{k/2}/\Gamma(k/2+1)$, which gives $\omega_0 = 1$, $\omega_1 = 2$, $\omega_2 = \pi$ and σ_0 and σ_1 are respectively the rms of the field and its first derivatives. These analytical expressions represent useful descriptions of the MFs which, for the case of a Gaussian random field, can be factorized as a function of the threshold and another of the shape and amplitude of the C_ℓ. We will use both the unnormalized (V_k) and unormalized (ν_k) MFs in the Gaussianity tests performed in this section. The unnormalized functionals are computed with a code that was used for the analysis of Archeops data (Curto et al. 2007) and has been thoroughly validated with *Planck* simulations, while for the normalized ones a code adapted to the high resolution *Planck* data and described in Ducout et al. (2012) is used.

By combining the MFs curves into a vector $\boldsymbol{y}$ of size $n = n_{\rm thresholds} \times n_{\rm functionals}$, a null hypothesis test can be performed using a χ^2 statistic given by:

$$\chi^2(\boldsymbol{y}) = (\boldsymbol{y} - \langle \boldsymbol{y}_{\rm G} \rangle)^{\rm T} \mathsf{C}^{-1} (\boldsymbol{y} - \langle \boldsymbol{y}_{\rm G} \rangle) \quad (21)$$

where $\boldsymbol{y}$ represents the MFs of the data, $\boldsymbol{y}_{\rm G}$ those of the simulations and C is the covariance matrix. In order to assure convergence, in the case of the four normalized MFs C is estimated from 10^4 Gaussian simulations, drawn from the *Planck* fiducial power spectrum, having the same instrumental properties of effective beam and noise as the data, the same applied mask and which have been processed in the same way to reach the corresponding resolution. For the three unnormalized MFs, C was estimated from only 10^3 FFP6 simulations that proved to be sufficient for convergence. We compare the $\chi^2_{\rm Planck}$ obtained from the data to the χ^2 obtained from those simulations, and report the probability of having a value of χ^2 larger than the measured one, $P\left(\chi^2 > \chi^2_{\rm Planck}\right)$. We explore different resolutions represented by the parameter $N_{\rm side}$, different methods of component separation (`Commander-Ruler`, `NILC`, `SEVEM`, and `SMICA`) and different sky coverages.

First, the three unnormalized MFs (V_k as a function of ν, $k = 0, 1, 2$) are used to construct a test of the null hypothesis. The test assesses not only the primordial Gaussian hypothesis, but also whether the data is correctly represented by the simulations in terms of power spectrum, systematics and the lensing effect. A set of 17 thresholds between -4 and $+4$ in steps of 0.5 are considered. The comparison between the MFs of the data provided by the four component separation methods and those corresponding to each of the four sets of 10^3 FFP6 simulations representing each method, for the standard U73 mask, are shown in Fig. 10. From that figure, a deviation at a level of $\approx 2\sigma$ can be seen for the contour and genus curves at a resolution $N_{\rm side} = 512$. The situation is very similar for the analyses performed at other resolutions, $N_{\rm side} = 1024$, 256 and 128. Although the deviation is not particularly compelling because of the correlations among neighbouring thresholds, it is worth

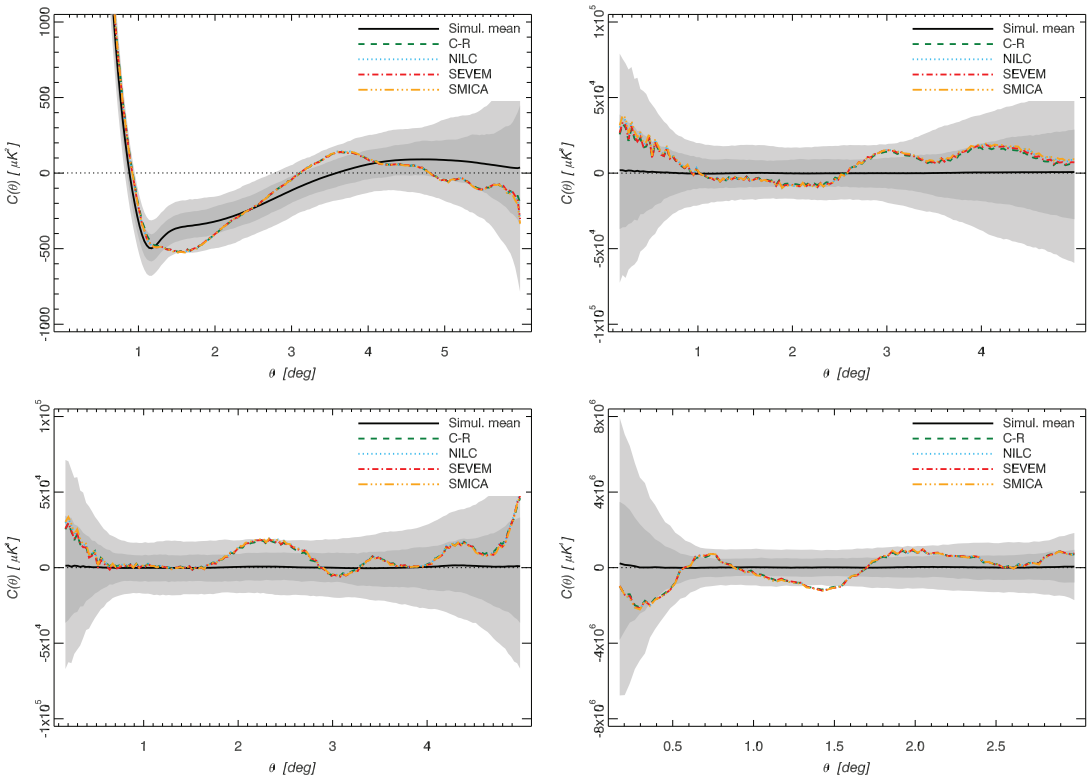


Fig. 9. The 2-point (upper left), pseudo-collapsed (upper right), equilateral (lower left) 3-point and reduced rhombic 4-point (lower right) functions averaged over the disc set for the $N_{\rm side} = 2048$ CMB estimates. Estimates of the multipoles for $\ell \leq 60$ are removed from the sky maps. The black solid line indicates the mean for 1000 MC simulations and the shaded dark and light grey regions indicate the 68% and 95% confidence regions, respectively.

mentioning that a possible explanation is the background of unresolved sources that has been detected in *Planck* data with the bispectrum estimators (see Planck Collaboration XXIV (2013)). In order to understand the effect of unresolved sources on the MFs, we added the point source residuals derived from the FFP6 simulations as processed by the SEVEM algorithm to 100 realisations which were then analysed. We conclude that the background of unresolved sources may be responsible for at least part of the excess signal that is detected. The corresponding probabilities $P\left(\chi^2 > \chi^2_{\rm Planck}\right)$ derived from the MF values for each of the four component separation methods and resolutions are given in Table 12. The full resolution maps have been degraded to the lower resolution ones following the procedure described in section 2. All the cases considered are compatible with the null hypothesis.

In the second case, the four normalized MFs $v_k = V_k/A_k$ ($k = 0, 1, 2, 3$) are used for the null hypothesis test. A set of 26 thresholds equally spaced between −3.5 and +3.5 are considered. The normalization factor A_k is estimated directly from the maps, having computed previously the moments σ_0 and σ_1. This normalization minimizes the dependence of the MFs on the power spectrum, thereby decreasing the cosmic variance and improving their sensitivity to deviations from Gaussianity. The resolutions considered in this case are $N_{side} = 2048, 1024, 512, 256$ and 128. For the highest resolution $N_{\rm side} = 2048$, the map is smoothed with a Gaussian smoothing kernel with a width $\theta_{\rm FWHM} = 5'$, in order to decrease the noise level. We use the standard U73 mask, inpainting the smallest point sources. The maps at lower resolution are constructed by the standard simple degrading process applied to the original map at $N_{\rm side} = 2048$, and the corresponding masks are degraded following a conservative procedure such that any degraded pixel with a value < 0.8 is set to zero (as explained in section 2). The results of the analysis

Table 12. Non-directional Gaussianity tests using unnormalized MFs: $P\left(\chi^2 > \chi^2_{\rm Planck}\right)$ as a function of sky resolution for different component separation methods.

$N_{\rm side}$	1024	512	256	128
C-R	0.812	0.299	0.482	0.357
NILC	0.993	0.567	0.354	0.234
SEVEM	0.925	0.911	0.738	0.094
SMICA	0.874	0.675	0.426	0.213

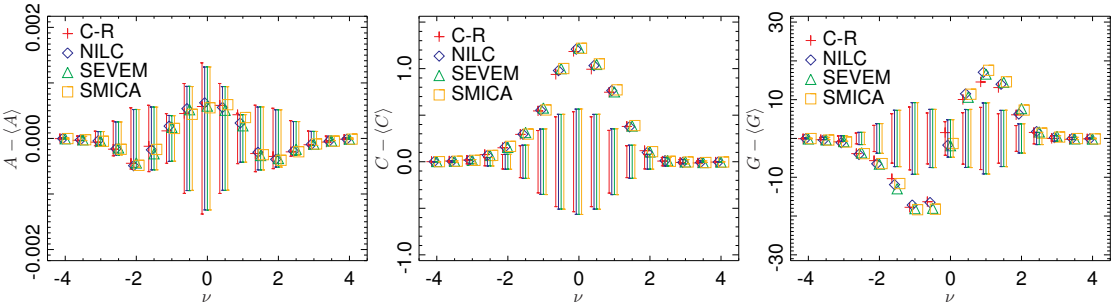


Fig. 10. Difference of the data MFs (unnormalized) with respect to the average of the curves obtained with realistic *Planck* simulations for several cleaned maps. From left to right: Area, Contour, Genus. The error-bars represent the 1σ (68%CL) dispersions around the mean obtained with simulations.

performed on the SMICA map at different resolutions are presented in Table 13. The results of the analysis performed on the different component separation methods at the highest resolution ($N_{\rm side} = 2048$) are presented in Table 14. The difference of the normalized MFs with respect to the expected values of the null hypothesis as a function of the threshold ν are shown in Fig. 11. A slight deviation in $N_{\rm cluster}(\nu)$ is noticeable at thresholds $\nu \approx 0$, however it is not very compelling since the values at neighboring thresholds are very correlated and this correlation is taken into account in the χ^2 statistics. Finally, we analyse the dependence of the normalized MFs on the sky coverage. We use the standard U73 mask and then decrease the sky coverage by using CL65, CL48 and CL25 masks in combination with a special point source mask that is based on the U73 mask. The fraction of sky left unmasked in the combined masks is 62%, 46% and 23%, respectively. The point source mask is inpainted previously to the analysis. The curves obtained for the different sky coverages are presented in Fig. 12, for the SMICA method. Results of the χ^2 analysis of the data as a function of sky coverage are compiled in Table 15. All the cases considered are compatible with the null hypothesis.

Table 13. Non-directional Gaussianity tests using normalized Minkowski Functionals: dependence of $P\left(\chi^2 > \chi^2_{\rm Planck}\right)$ on Sky resolution.

$N_{\rm side}$	2048	1024	512	256	128
Normalized MFs	0.358	0.356	0.245	0.225	0.223

Table 14. Non directional Gaussianity tests using normalized Minkowski Functionals: Dependence on omponent separation methods.

Method	C-R	NILC	SEVEM	SMICA
$P\left(\chi^2 > \chi^2_{\rm Planck}\right)$	0.288	0.303	0.415	0.358

In summary, we find that the data are globally consistent with the primordial Gaussian hypothesis, and no strong deviation is found between the data and realistic simulations for both the unnormalized and normalized MFs. We would like to remark that a certain level of non-Gaussianity is expected from lensing and, in particular, from the ISW-lensing signal, thus it is important to compare the data to realistic *lensed* simulations.

Table 15. Non directional Gaussianity tests using normalized Minkowski Functionals : Sky coverage.

$f_{\rm sky}$	0.73	0.62	0.46	0.23
$P\left(\chi^2 > \chi^2_{\rm Planck}\right)$	0.358	0.042	0.670	0.780

4.5. Wavelet statistics

A broad range of wavelets have been used in the analysis of CMB data, but in this paper we consider the Spherical Mexican Hat wavelet (SMHW, Martínez-González et al. 2002).

The SMHW is an example of a continuous, non-orthogonal wavelet. Given a signal on the sky, $T(p)$, where p represents a given position/pixel wich is a function of the co-latitude θ and longitude ϕ (also defined by the unit direction vector $\mathbf{x}$), the SMHW coefficients at a given scale R, $\omega_T(R, p)$, are obtained by convolution:

$$\omega_T(R, p) = \sum_{\ell=0}^{\ell_{\rm max}} \sum_{m=-\ell}^{m=\ell} t_{\ell m} W_\ell^{\rm SMHW}(R)\, Y_{\ell m}(p)\,, \qquad (22)$$

where $W_\ell^{\rm SMHW}(R)$ is the window function associated with the SMHW, $\ell_{\rm max}$ is the maximum multipole allowed by the corresponding HEALPix pixelization, $Y_{\ell m}(p)$ is the spherical harmonic basis, and $t_{\ell m}$ are the spherical harmonic coefficients of the analysed map:

$$t_{\ell m} = \int {\rm d}\Omega Y^*_{\ell m}(p)\, T(p)\,, \qquad (23)$$

where ${\rm d}\Omega = {\rm d}\theta \sin\theta {\rm d}\phi$ and * denotes complex conjugation.

Several statistics can be computed from the wavelet coefficients map, in particular, the first moments: the dispersion σ_R, the skewness S_R, and the kurtosis K_R (as a function of scale R).

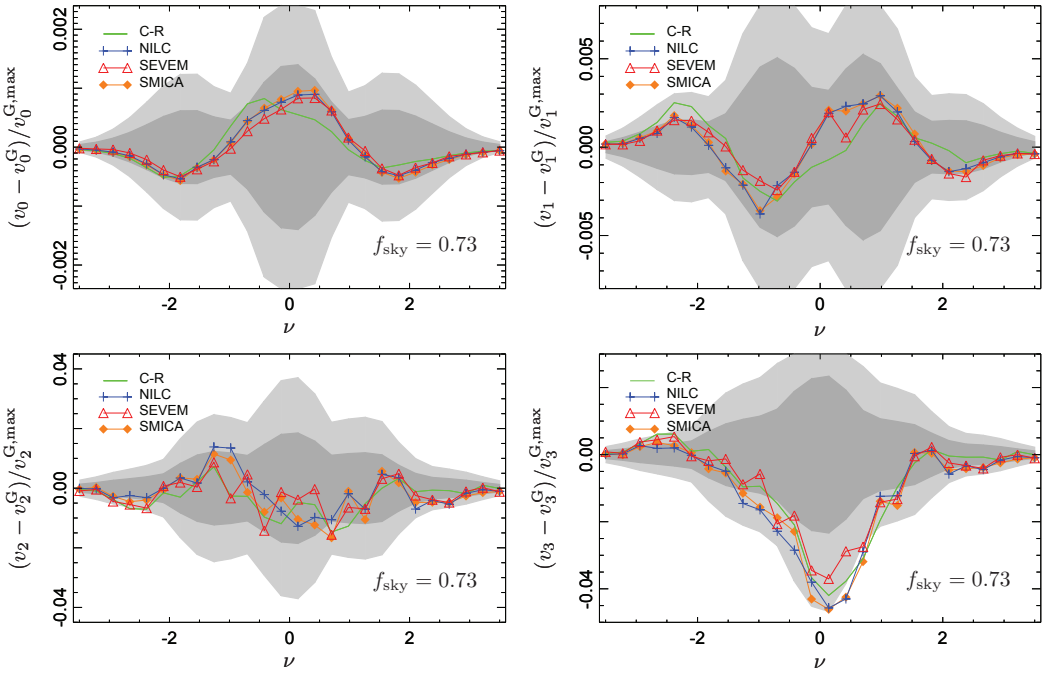


Fig. 11. Difference of the normalized MFs obtained from the data with respect to the expected values of the null hypothesis, for the different component separation methods. From left to right and top to bottom: Area, Contour, Genus and $N_{\rm cluster}$. The grey bands represent the 1 and 2σ dispersions around zero, based on realistic *Planck* simulations including lensing, for C-R method.

It is interesting to notice that in the case of Gaussian temperature fluctuations the linear transformation involved in the determination of the wavelet coefficients (eqs. 22,23) guarantees that Gaussianity is preserved.

The study of the moments of the distribution of the CMB temperature fluctuations, as a function of the scale, is a standard approach to test the null hypothesis. We have performed a full resolution multi-scale analysis of the four CMB clean maps and computed the quantities $\sigma_{\rm R}$, $S_{\rm R}$ and $K_{\rm R}$ from the SMHW wavelet coefficients at 18 scales, R = {2, 4, 7, 14, 25, 50, 75, 100, 150, 200, 250, 300, 400, 500, 600, 750, 900, 1050}, in arc-minutes. These are compared to the standard *Planck* simulations.

As explained in Vielva et al. (2004), when computing the SMHW coefficients of a masked data set, artefacts are introduced close to the mask that degrade the performance of any null hypothesis tests. We therefore define a set of *exclusion masks* such that, at each scale, an extra region of the sky is excluded when performing any statistical test. The exclusion mask for a given scale R is defined as follows: we build an auxiliary mask by removing from the U73 mask all the features associated with compact objects, and degrade this auxiliary mask to $N_{\rm side} = 1024$ (imposing a restrictive cut); a first temporary mask is obtained by extending the borders of this auxiliary mask by a distance of twice R; a second temporary mask is obtained, first, by convolving the auxiliary mask with the SMHW at that particular scale R and, second, by imposing that any pixel of that second temporary mask with an absolute value lower than 0.1 is masked, whereas the remaining ones are set to 1; the two temporary masks are multiplied to yield a single mask that is upgraded to $N_{\rm side} = 2048$; finally, the final exclusion mask is obtained by multiplying this mask by the parent U73 mask.

The comparison of the four CMB maps with the corresponding simulations is summarized in Fig. 13. The three panels show (from left to right) the statistical significance of the standard deviation, the skewness and the kurtosis (as a function of the SMHW scale). The points represent the upper tail probability associated to a given statistic, i.e., the fraction of the simulations that present a value of a given statistic equal to or greater than the one obtained for the data. In fact, we define a modified upper tail probability: if an upper tail probability is larger than 0.5, then a new quantity is defined as 1 minus that upper tail probability. Hence, this modified definition of the upper tail probability is constrained between 10^{-3} (the minimum value that can be imposed with 1 000 simulations) and 0.5. Overall, the agreement between the four CMB maps is quite good, showing that all of them provide a consistent estimation of the true CMB. However, several aspects need to be discussed. Let us clarify that the differences among the CMB methods for small modified upper tail probabilities are expected to be larger than for large modified upper tail probabilities. This is because a small modified upper tail probability is determined by a small number of simulations and, therefore, has a relatively large error bar. In other words, the tails of the distributions of the different statistics are quite sparsely sampled.

We will distinguish between the small ($R \lesssim 10'$), intermediate ($10' \lesssim R \lesssim 500'$) and the large ($R \gtrsim 500'$) scale regimes. Let us focus on the three statistics independently. We will highlight

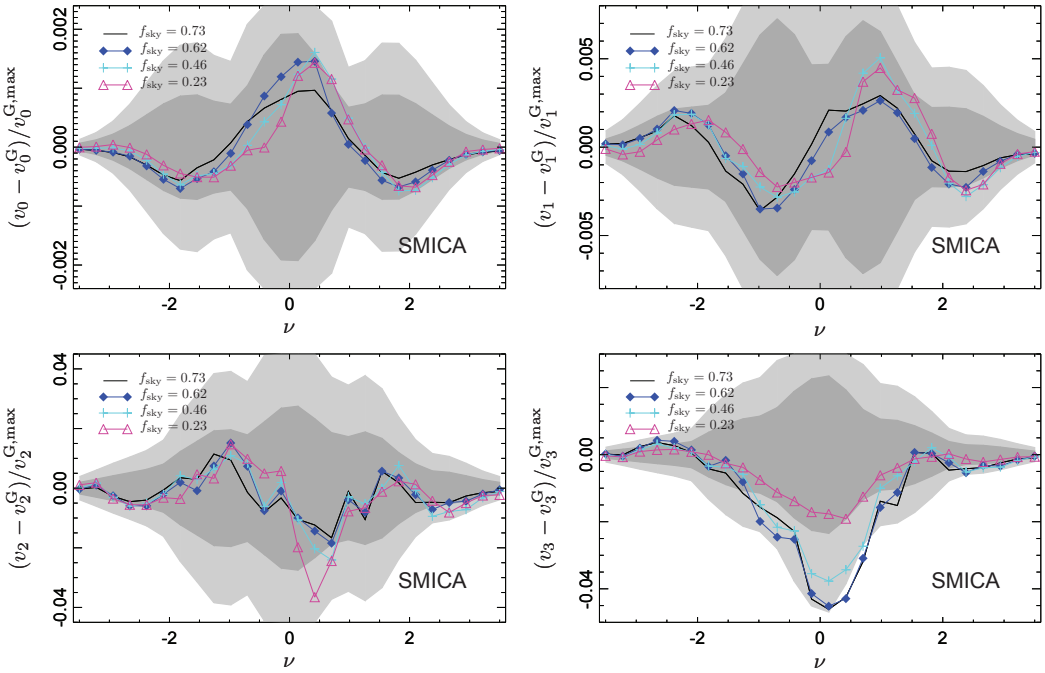


Fig. 12. Difference of the normalized MFs obtained from the data with respect to the expected value of the null hypothesis for several sky coverages. The SMICA map is considered. From left to right and top to bottom: Area, Contour, Genus and $N_{\rm cluster}$. The grey bands represent the 1 and 2σ dispersions around zero, based on realistic *Planck* simulations including lensing, for $f_{\rm sky} = 0.23$.

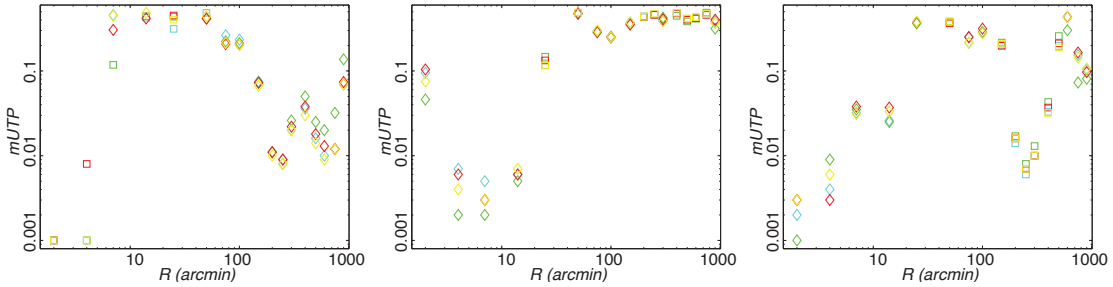


Fig. 13. Standard deviation (left), skewness (centre) and kurtosis (right) of the SMHW coefficients as a function of the wavelet scale R. Results are given for the four *Planck* CMB maps (green: Commander-Ruler, light-blue: NILC; red: SEVEM; yellow: SMICA). Modified upper tail probabilities (*mUTP*, see text for details) are obtained by comparing with 1000 simulations processed through the component separation pipelines. Squares represent modified upper tail probabilities that correspond to an actual upper tail probability above 0.5; diamonds represent upper tail probabilities below 0.5.

the most important features and, afterwards, we will try to find an explanation for them:

- On the smallest scales, the four CMB maps show a dispersion in SMHW coefficients significantly larger than seen in the simulations. However on larger scales, the dispersion is systematically below the median of the simulations and, on scales of $R \approx 5°$, the modified upper tail probability is approximately 0.015.
- Regarding the skewness, all four maps yield a value that is significantly lower (with a modified upper tail probability of around 0.004) than expected from the simulations in the small scale regime (except for the smallest one, where the deviation is around 0.07). The rest of the scales are fairly compatible with the null hypothesis.
- The kurtosis is also smaller than expected in the small scale regime. Overall, the modified upper tail probability is about 0.03. At scales of around $300'$, an anomalously large value

(modified upper tail probability of approximately 0.01) is found.

These results are compatible with the values reported for *WMAP* data (Vielva et al. 2004; Cruz et al. 2005), over the scales common to both experiments (i.e., $R > 10'$). In particular, the large value of the kurtosis has been associated with the Cold Spot (Vielva et al. 2004). We wil return to this topic specifically in Sect. 5.8. The low variance of the wavelet coefficients was previously seen in Vielva et al. (2004); Wiaux et al. (2008). In addition, the low dispersion at scales above a few degrees is likely to be related to the low variance anomaly detected in *WMAP* (Monteserín et al. 2008; Cruz et al. 2011), that is also seen in the *Planck* data (see Sect. 4.1).

We have also studied the robustness of the results for different masking scenarios. In particular, we have investigated variations in the results when we adopt, as auxiliary masks to define the exclusion masks, the two CG70 and CG60 masks removing 30% and 40% of the sky, respectively. Note that the auxiliary masks obtained from the U73 mask already cut around 20% of the sky. The corresponding results for the SMICA map are presented in Fig. 14. The conclusions are similar for the other CMB maps. For the dispersion of the wavelet coefficients, we do not notice any significant change in the anomalously high value obtained for the SMICA map at the smallest scales However, some changes are observed at larger scales. In this regime, it seems that the most significant deviation occurs for the CG70 mask (modified upper tail probability of around 0.005), whereas similar (and slightly less significant) modified upper tail probabilities are obtained for both the U73 (modified upper tail probability of approximately0.015) and the CG60 (modified upper tail probability of about 0.01) masks. A possible explanation for this behaviour would be that a less restrictive mask admits some residual contamination from Galactic foregrounds, thus increasing the dispersion of the wavelet coefficients, and artificially increasing their inconsistency with the null hypothesis. In principle, the larger the Galactic cut, the lower would be the dispersion of the wavelet coefficients (assuming that some residual contamination of the Galactic foregrounds is left) and, therefore, the smaller the upper tail probability. However, as we already said, the modified upper tail probability for the dispersion is higher for the CG60 mask than for the CG70 mask. This apparent contradiction could be resolved by accounting for the larger sampling variance for smaller areas, that would result in a lower significance for the anomaly.

The anomalous kurtosis at scales of $R \approx 300'$ shows an overall stable modified upper tail probability of around $0.01 - -0.03$. In the small scale regime, the differences are better defined: the smaller the mask, the more significant the deviation (characterized by the low value of the kurtosis). In particular, the modified upper tail probability associated with the CG60 mask is 0.001, around 0.009 for CG70, and approximately 0.03 for the U73 mask. A similar pattern is also observed for the skewness on these scales, although the three masks results in more similar upper tail probabilities, between around 0.001 and 0.007 (except for the smallest scale).

It is therefore clear that there is some inconsistency between the CMB data and the corresponding simulations. On intermediate scales, both the low dispersion and the high kurtosis could be related to previously known anomalies: the low variance and the Cold Spot. On the smallest scales, the three statistics report a low upper tail probability independently of the mask coverage — it is important to determine what this inconsistency is due to. Besides the possibility that it is an intrinsic cosmological signal, the non-Gaussianity could be caused either by instrumental systematics or residual foreground contamination.

In the former case, we have considered whether the origin of the signal could be related to properties of the noise that are inadequately modelled by the simulations. In particular, we have studied the statistical properties of the half-ring half-difference maps generated by the four component separation algorithms as proxies for the the noise present in the CMB maps. Although in detail there are some discrepancies between these noise estimates and the simulated ones, they are not compatible with the inconsistencies observed between the CMB map and simulations. Therefore, a systematic effect associated with the instrumental noise does not provide a satisfactory explanation for the small-scale deviations.

In the latter case, an obvious candidate is due to the contribution from residual unresolved point sources in the clean CMB maps. Although the brightest point sources are masked, and the component separation process itself can suppress the amplitude of the unresolved background of point sources, some signal will remain. Indeed, in Planck Collaboration XXIV (2013) it has been determined that the bispectrum of this contribution is clearly detected in the four CMB *Planck* maps, at a significance in excess of 4σ. In addition, the dispersion of the wavelet coefficients is higher than expected, which is also compatible with the presence of an additional signal. We therefore consider this as the most likely non-CMB explanation for the observed signal.

4.6. Bispectrum

The CMB bispectrum is the three point correlator of the $a_{\ell m}$ coefficients,

$$B^{\ell_1\ell_2\ell_3}_{m_1m_2m_3} = a_{\ell_1m_1}a_{\ell_2m_2}a_{\ell_3m_3}\,. \tag{24}$$

In this paper, we focus on the bispectrum reconstruction as a blind test of non-Gaussianity. Therefore, we assume we are seeking a non-trivial bispectrum that has arisen through a physical process which is statistically isotropic, that is, we can employ the angle-averaged bispectrum $B_{\ell_1\ell_2\ell_3}$,

$$B_{\ell_1\ell_2\ell_3} = \sum_{m_i} h^{-1}_{\ell_1\ell_2\ell_3}\mathcal{G}^{\ell_1\ell_2\ell_3}_{m_1m_2m_3}B^{\ell_1\ell_2\ell_3}_{m_1m_2m_3}\,, \tag{25}$$

where $h_{\ell_1\ell_2\ell_3}$ is a geometrical factor,

$$h_{\ell_1\ell_2\ell_3} = \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi}}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ 0 & 0 & 0\end{pmatrix}, \tag{26}$$

and $\mathcal{G}^{\ell_1\,\ell_2\,\ell_3}_{m_1m_2m_3}$ is the Gaunt integral,

$$\begin{aligned}\mathcal{G}^{\ell_1\ell_2\ell_3}_{m_1m_2m_3} &\equiv \int d\Omega\, Y_{\ell_1m_1}(\boldsymbol{n})\, Y_{\ell_2m_2}(\boldsymbol{n})\, Y_{\ell_3m_3}(\boldsymbol{n})\\ &= h_{\ell_1\ell_2\ell_3}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix},\end{aligned} \tag{27}$$

with the usual Wigner-3j symbol $\left(\begin{smallmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{smallmatrix}\right)$. It is more convenient to eliminate the geometrical factors entirely and to work with the reduced bispectrum which is defined as

$$b_{\ell_1\ell_2\ell_3} = h^{-1}_{\ell_1\ell_2\ell_3}B_{\ell_1\ell_2\ell_3}\,. \tag{28}$$

Note that the CMB bispectrum $b_{\ell_1\ell_2\ell_3}$ is defined on a tetrahedral domain of multipole triples $\{\ell_1\ell_2\ell_3\}$ satisfying both a triangle condition and a limit given by the maximum resolution $\ell_{\rm max}$ of the experiment. A much more extensive introduction to the bispectrum can be found in Planck Collaboration XXIV (2013).

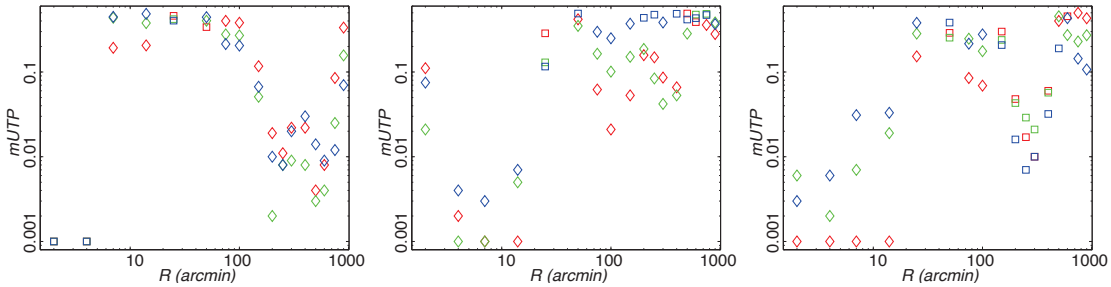


Fig. 14. Standard deviation (left), skewness (centre) and kurtosis (right) of the SMHW coefficients as a function of the wavelet scale R. Results are given for the SMICA CMB map. Several masking scenarios are compared: red: CG60 mask (cutting out 40% of the sky); green: CG70 mask (cutting out 30% of the sky); blue: U73 mask. The modified upper tail probabilities ($mUTP$) are defined in the text.

Modal, wavelet and binned bispectrum estimators filter the CMB map with separable basis functions

$$Q_{ijk}(\ell_1, \ell_2, \ell_3) = q_i(\ell_1)\, q_j(\ell_2)\, q_k(\ell_3) + \text{perms}\,, \tag{29}$$

to find the corresponding modal coefficients β_{ijk} (or β_n because it is convenient to order the ijk with label n). For appropriately orthonormalised basis functions $Q_{ijk}(\ell_1, \ell_2, \ell_3)$, these coefficients can be used to reconstruct the CMB bispectrum through the signal-to-noise weighted expansion

$$\frac{b_{\ell_1\ell_2\ell_3}}{\sqrt{C_{\ell_1}C_{\ell_2}C_{\ell_3}}} = \sum_n \beta_{ijk}\, Q_{ijk}(\ell_1, \ell_2, \ell_3)\,. \tag{30}$$

This reconstruction method has been extensively validated, showing the accurate recovery of CMB bispectra from non-Gaussian simulated maps, and it has been applied to the *WMAP* seven year data to reconstruct the full 3D CMB bispectrum (Fergusson et al. 2010b). To quantify whether or not there is a model-independent deviation from Gaussianity, we can consider the total integrated bispectrum. By summing over all multipoles, we can define an integrated nonlinearity parameter $\bar{F}_{\rm NL}$ which, with the orthonormal modal decomposition (30), becomes (Fergusson et al. 2010b)

$$\bar{F}^2_{\rm NL} = \frac{1}{N^2_{\rm loc}} \sum_{\ell_i} \frac{h^2_{\ell_1\ell_2\ell_3} b^2_{\ell_1\ell_2\ell_3}}{C_{\ell_1}C_{\ell_2}C_{\ell_3}} = \frac{\sum_{ijk}\beta_{ijk}{}^2}{\sum_{ijk}\alpha^{\rm loc\,2}_{ijk}}\,. \tag{31}$$

where $N_{\rm loc}$ is the normalisation for the local $f_{\rm NL} = 1$ model (with coefficients $\alpha^{\rm loc}_{ink}$). For ideal Gaussian CMB maps, the quantity $\bar{F}^2_{\rm NL}$ should obey a χ^2-distribution with a mean given by the number of degrees of freedom (the modes) $\mu = n_{\rm max}$ and with a variance $\sigma^2 = 2n_{\rm max}$. Assuming that the three-point correlator is the leading non-Gaussian contribution, then $\bar{F}_{\rm NL}$ provides a blind test for the presence of any integrated CMB bispectrum (once the expected two-point term is subtracted). We note that his is less sensitive than targeted searches for particular bispectrum shapes.

First, we discuss reconstructions from the modal estimator which has passed successfully through the full suite of non-Gaussian bispectrum validation tests (for further details about bispectrum estimators, see Planck Collaboration XXIV 2013). We have applied this to the *Planck* temperature maps for the foreground-separation techniques NILC SEVEM and SMICA, using two alternative sets of hybrid basis functions in order to cross-check results and identify particular signals. These are

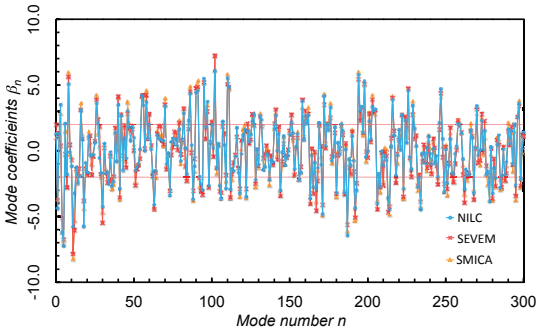


Fig. 15. *Planck* recovered bispectrum coefficients β^R_n for the mode expansion (30) using hybrid Fourier modes (augmented with local and ISW modes). There is remarkable consistency between results from the different component separation methods, NILC, SEVEM, and SMICA. The variance from simulated noise maps is nearly constant for each of the 300 modes, with the average $\pm 1\sigma$ variation shown in red.

Fourier modes ($n_{\rm max} = 300$) augmented a local SW mode and the separable ISW modes and a hybrid polynomial/local basis with $n_{\rm max} = 600$, previously described in Fergusson et al. (2010a). These basis function sets ensured excellent correlation with primordial modes and the ISW signal. As with all the other bispectrum analyses based on spherical harmonic coefficients, we used the U73 mask to which we applied inpainting. Together with the foreground separated maps, noise simulations were provided which were used to calibrate the estimator's linear correction term and to determine the variance.

The modal coefficients β^R_n extracted from the *Planck* NILC, SEVEM, and SMICA maps are shown in Fig. 15 for the hybrid Fourier basis. These amplitudes show remarkable consistency between the different maps, with shape cross-correlations better than 96% and the overall amplitudes to within 7% agreement. This demonstrates that the indendent foreground separation techniques do not appear to be introducing spurious non-Gaussianity. The β^R_n coefficients have a roughly constant variance, so anomalously large modes can be easily identified. For

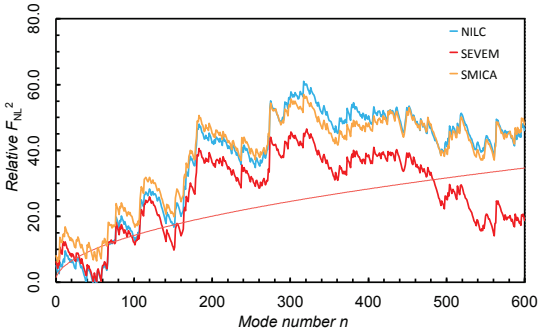


Fig. 16. Cumulative sum of orthonormal mode contributions β_n^{R2} to the total integrated bispectrum $\bar{F}^2_{\rm NL}$ defined in (31) . The relative quantity $\bar{F}^2_{\rm NL} = F^2_{\rm NL} - F^{\rm G\,2}_{\rm NL}$ is plotted, where $F^{\rm G\,2}_{\rm NL}$ is the mean obtained from 200 CMB Gaussian maps and the standard deviation is the red line. A hybrid polynomial basis $n_{\rm max} = 600$ is employed in the signal-dominated region $\ell < 1500$. This χ^2-test for the independent modes is cumulatively consistent with Gaussianity.

example, we have subtracted the expected ISW signal and the estimated point source contributions, explaining the large signal at low n. The corresponding quantity $\bar{F}^2_{\rm NL}$ defined in (31), that can be seen in Fig. 16, shows consistency with the null hypothesis. Using the modal expansion (30), we have reconstructed the full 3D *Planck* bispectrum which is illustrated in Fig. 17 for SMICA (large) but also NILC and SEVEM; the reconstructions are visually indistinguishable. There are some striking features evident, notably the presence of a significant ISW modal contribution in the squeezed limit along the edges of the tetrapyd which has an oscillatory and flattened appearance. At large multipoles ℓ approaching $\ell_{\rm max} = 2000$, there is increased randomness in the reconstruction due to the rise in experimental noise and some evidence for a residual point source contribution. For the present *Planck* estimator configurations, the modal bispectrum estimator is more democratic, that is, it is capable of resolving the large-ℓ contributions near $\ell_{\rm max}$ seen in Fig. 15, and not only the multipoles associated with primordial models.

In Fig. 18, we show a comparison of the $\ell < 500$ *Planck* bispectrum signal and that reconstructed from the *WMAP* seven-year data (Fergusson et al. 2010b). Here for consistency we show the *Planck* signal from the second polynomial basis, since polynomials were used in the original *WMAP*7 analysis. The *Planck* signal pattern correlates well with the *WMAP* bispectrum obtained previously, despite the different domains used for the modal analysis of the two different experiments.

Similarly to the modal bispectrum, a wavelet decomposition can be used to reconstruct the bispectrum. Here we use the continuous, non-orthogonal Spherical Mexican Hat Wavelet (SMHW, Martínez-González et al. 2002). Cubic moments q_{ijk} are defined in terms of the SMHW coefficients for three different angular scales R_i, R_j, R_k (Curto et al. 2009b,a, 2010, 2011a,b)

$$q_{ijk} = \frac{1}{4\pi}\frac{1}{\sigma_i\sigma_j\sigma_k}\int d\boldsymbol{n} w(R_i,\boldsymbol{n}) w(R_j,\boldsymbol{n}) w(R_k,\boldsymbol{n}) \qquad (32)$$

where σ_i is the dispersion of the wavelet coefficient map $w(R_i, \boldsymbol{n})$. Considering the covariance matrix of the q_{ijk} moments, $\mathsf{C} \equiv \langle \mathsf{q}\mathsf{q}^{\rm T}\rangle$, and its eigenvector decomposition, $\mathsf{C} = \mathsf{R}\mathsf{D}\mathsf{R}^{\rm T}$, with R the eigenvector matrix and D the eigenvalue matrix, a new set of quantities $\boldsymbol{y} \equiv \mathsf{D}^{1/2}\mathsf{R}^{\rm T}\mathsf{q}$ is defined. Considering the decorrelation produced by the convolution of the SMHW on the temperature anisotropies and applying the central limit theorem to the averages defined in Eq. 32, then the q_{ijk} quantities are expected to have a nearly Gaussian distribution. Therefore, the $\boldsymbol{y}$ quantities are nearly multinormal and satisfy $\langle \boldsymbol{y}\boldsymbol{y}^{\rm T}\rangle = \mathsf{I}$ and $\langle \boldsymbol{y}\rangle = \boldsymbol{0}$ (Curto et al. 2011a).

We have computed this reconstruction using the *Planck* data and compared with the null hypothesis (Gaussian *Planck* simulations). The considered data map is the resulting map after foreground cleaning based on different cleaning procedures: Commander-Ruler, NILC, SEVEM, and SMICA. The mask used is the U73 one (contrary to the modal reconstruction, no inpainting of the point sources is made in this case). In Fig. 19 the $\mathbf{y}$ statistics corresponding to the *Planck* data are plotted and compared with the 3σ error-bars obtained with *Planck* Gaussian simulations. From the list of different q_{ijk} statistics corresponding to the 16 angular scales described in Planck Collaboration XXIV (2013), there are 11, 4, 3, 3 statistics with $|y_i| \geq 3$ (corresponding to Commander-Ruler, NILC, SEVEM, and SMICA respectively). The error-bars are obtained with *Planck* simulations for each type of component separation cleaned map. The error-bars of the y_i statistics for low indices i are associated to large scales where the q statistics have a less Gaussian-like shape. The $\boldsymbol{y}$ statistics are combined into a χ^2 test after a principal component analysis with a threshold of 10^{12} (Curto et al. 2011a) and compared with the χ^2 statistics obtained from *Planck* Gaussian simulations for each type of component separation method (see Table 16). The χ^2 statistic corresponding to the data is compatible with the values obtained from Gaussian simulations according to the cumulative probability $P(\chi^2 \geq \chi^2_{\rm data})$, as can be seen in Table 16. Therefore the wavelet bispectrum reconstruction does not detect a significant amplitude of bispectrum in the considered data maps. Details on the constraints on the amplitude of different bispectrum shapes are presented in Planck Collaboration XXIV (2013).

Table 16. χ^2 statistics based on the wavelet bispectrum reconstruction y_i statistics for the foreground cleaneddatadata map. Considered data map: combined map cleaned with C-R, NILC, SEVEM, and SMICA.

Method	$\chi^2_{\rm data}$	DOF	$\langle\chi^2\rangle$	σ	$P(\chi^2 \geq \chi^2_{\rm data})$
C-R	874	690	740	87	0.074
NILC	883	682	731	83	0.045
SEVEM	858	682	731	83	0.070
SMICA	878	682	732	83	0.058

5. Intriguing inconsistencies – *WMAP* anomalies revisited.

In the previous section, we have established that the *Planck* data shows little evidence for non-Gaussianity beyond that expected due to the ISW-lensing effect (which is accounted for directly by simulations), and contributions from residual unresolved point sources. The exceptions are on large-angular scales where features consistent with various anomalies previously seen in the *WMAP* data have been observed. In this section, we explicitly consider several of the most important anomalies detected

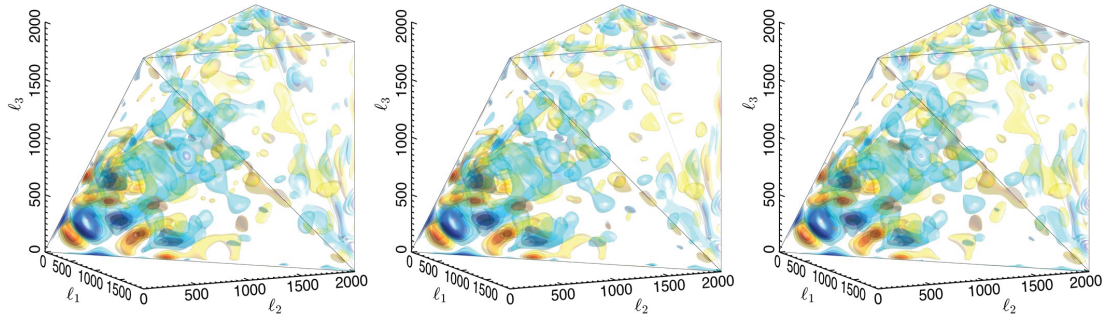


Fig. 17. Full 3D CMB bispectrum recovered from the Planck foreground-separated maps (modes illustrated in fig. 15, including SMICA (left), NILC (centre) and SEVEM (right). These are plotted in three-dimensions with multipole coordinates $\{\ell_1, \ell_2, \ell_3\}$; the triangle condition restricts the bispectrum to a tetrahedral domain out to the experimental resolution limit $\ell < \ell_{\rm max} = 2000$. Several density contours are plotted with red positive and blue negative. The bispectra from different component-separation methods are almost indistinguishable with the same features also appearing in Fourier and polynomial expansions. Note the central and flattened features for $\ell < 1200$ and also the oscillating CMB ISW lensing signal in the squeezed limit along the edges of the tetrapyd.

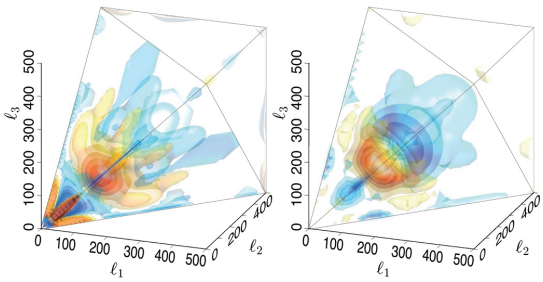


Fig. 18. Comparison between the *WMAP* seven-year bispectrum signal (left) (Fergusson et al. (2010b)) and the low-ℓ signal of Planck (right) reconstructed from the SMICA foreground-separated map (in both cases using polynomial modes). The same basic patterns are observed in both bispectra, including an apparent central 'oscillatory' feature.

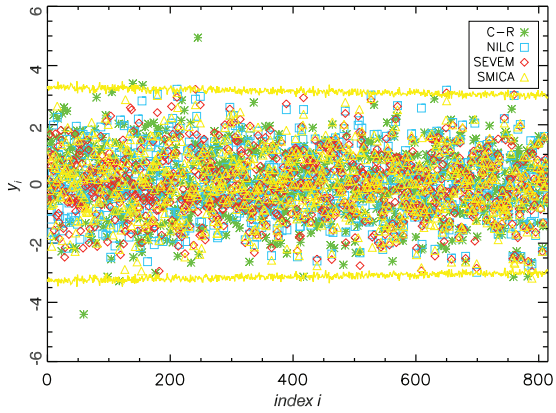


Fig. 19. The wavelet bispectrum reconstruction y_i statistics for the foreground cleaned *Planck* data map. Considered data map: combined map cleaned with C−R, NILC, SEVEM and SMICA. The solid yellow lines represent the 3σ error-bars for SMICA (similar error-bars are obtained for C−R, NILC, and SEVEM maps).

in the *WMAP* data, namely the quadrupole-octopole alignment (Sect. 5.1), the low variance (Sect. 5.2), hemispherical asymmetry (Sect. 5.3), phase correlations (Sect. 5.4), dipolar power modulation (Sect. 5.5), generalized power modulation (Sect. 5.6), parity asymmetry (Sect. 5.7) and the Cold Spot (Sect. 5.8). Each of these anomalies may represent different violations of the fundamental properties of isotropy and/or Gaussianity of the CMB data which are assumed in the estimation of the CMB power spectrum.

There is an ongoing debate about the significance of these anomalies in the literature. A critical issue relates to the role of a posteriori choices — whether interesting features in the data bias the choice of statistical test or if arbitrary choices in the subsequent data analysis enhance the significance of the features. Indeed, the *WMAP* team (Bennett et al. 2011) contends that the anomalies are significantly over-interpreted due to such selections, whilst other authors claim highly significant and robust detections. Therefore, care must be taken to address the issue, since our analyses are necessarily follow up tests of the previous *WMAP* investigations. However a careful and fair statistical treatment can allow us to study possible links among the anomalies and to search for a physical interpretation.

5.1. Mode alignment

Tegmark et al. (2003) first reported a significant alignment between the orientation of the quadrupole and the octopole in the *WMAP* first year temperature data. We study this quadrupole-octopole alignment in the *Planck* data using the maximization of the angular momentum dispersion as described in de Oliveira-Costa et al. (2004). Specifically, we determine the orientation of the multipoles by finding the axis $\boldsymbol{n}$ around which the angular momentum dispersion

$$\sum_m m^2 |a_{\ell m}(\boldsymbol{n})|^2 \tag{33}$$

is maximized. Here, $a_{\ell m}(\boldsymbol{n})$ denotes the spherical harmonics coefficients of the CMB map in a rotated coordinate system with its z-axis in the $\boldsymbol{n}$-direction. This definition of the multipole-orientation has been devised for planar multipoles and is sim-

ply the direction perpendicular to the plane in which most of the power of the multipole lies. It is thus intuitive and easy to use. Note that the value of the statistic in Eq. (33) is the same for $-\boldsymbol{n}$ as for $\boldsymbol{n}$, i.e. the multipole orientation is defined only up to a sign.

An alternative method, based on the multipole vector decomposition (Copi et al. 2004; Schwarz et al. 2004; Bielewicz et al. 2005; Bielewicz & Riazuelo 2009) of the data has also been used to verify the robustness of the results presented here, and excellent consistency is found.

Residual foregrounds (mostly on the Galactic plane) present in the four *Planck* CMB estimates could influence the reconstruction of the low-order multipoles. However, when a mask applied, the resulting mode-coupling can also affect the reconstruction of the low-ℓ multipoles. We therefore utilise Wiener filtered maps computed from the data to which the U73 mask is applied. Specifically, we utilise the same implementation of the Wiener filter as used in Planck Collaboration P09A (2013) i.e., a messenger method as described by Elsner & Wandelt (2013). A direct inversion method for masked data (Efstathiou 2004; Bielewicz et al. 2004, 2013) is a possible alternative, but the Wiener filtered maps result in a significantly smaller uncertainty in the reconstructed orientation of the multipoles.

We then search for the preferred orientation by explicitly rotating the CMB map such that the z-axis pierces the centre of all the low resolution pixels defined at $N_{\rm side} = 16$, and then subsequently refine the search by using an $N_{\rm side} = 2048$ map. The angular resolution for the orientation of the multipoles is thus given by the distance between the pixel centers of the $N_{\rm side} = 2048$ map, which is of order $1.94'$. Figure 20 shows the Wiener filtered SMICA CMB sky, with the corresponding reconstruction of the quadrupole and octopole moments. The reconstructed orientations are quite robust with respect to the component separation method used for reconstructing the CMB. The significance of the alignment between the quadrupole and the octopole is assessed from the scalar product of their orientations, compared to values derived from the standard set of 1000 MC simulations. The orientation, the angular distance the scalar-product between quadrupole and octopole, and the probability of at least such an alignment to occur in an isotropic universe are summarised in Table 17 for each CMB map.

We find that, depending on the component separation method, the quadrupole and octopole orientations are misaligned by an amount between 9° and 13°. This is larger than the 3° reported recently by Bennett et al. (2012) for the 9-year *WMAP* ILC map. In consequence, our significance of the quadrupole-octopole alignment is substantially smaller than for the *WMAP* data, falling to almost 98% confidence level for the Commander-Ruler and SEVEM maps and 96.7% confidence level for the NILC map.

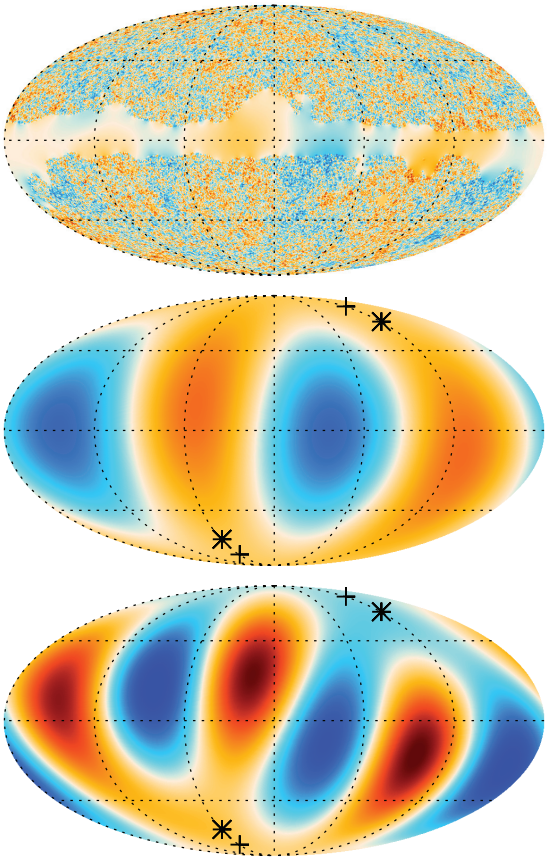


Fig. 20. Upper: The Wiener filtered SMICA CMB sky (temperature range ± 400 μK). Middle: the derived quadrupole (temperature range ± 35 μK). Lower: the derived octopole (temperature range ± 35 μK). Cross and star signs indicate axes of the quadrupole and octopole, respectively, around which the angular momentum dispersion is maximized.

5.2. Variance, skewness and kurtosis anomalies

A low value for the variance on the CMB sky was previously observed in the *WMAP* data by Monteserín et al. (2008) and Cruz et al. (2011), and confirmed for *Planck* in Sect. 4.1. Furthermore, the effect has also been seen in the wavelet analysis of Sect. 4.5 where the variance of the SMHW coefficients is low at scales between 400 and 600 arcmin (Fig. 13). In addition, anomalous behaviour was also observed for the skewness and kurtosis in low resolution maps at $N_{\rm side} = 16$. Here, we reassess these results and determine their robustness to masking and data selection. The former will allow us to determine whether a particular region is causing the anomalous behaviour, whilst the latter can establish whether foreground residuals could be responsible.

Table 18 and Fig. 21 present the results for the variance, skewness and kurtosis determined from the four CMB maps with the U73, CL58 and CL37 masks applied. Results are also computed for data within the ecliptic hemispheres surviving the U73 mask. The variance is low in all cases, with only small differences in significance observed for the different maps. Interestingly, the low variance seems to be localised in the northern ecliptic hemisphere. Conversely, anomalous values for the skewness and kurtosis are only apparent for the southern ecliptic hemisphere.

Since these results might be indicative of the presence of Galactic foreground residuals near the Galactic plane, we analyse the frequency dependence of the statistics as summarised in Table 19 and Fig. 22. The variance shows little frequency dependence for the considered masks and regions, whereas the skewness and kurtosis show a moderate frequency dependence when the U73 mask is applied, as also seen for the N-pdf in Sect. 4.2.

Table 17. Orientation of the low multipoles extracted from the different component separated CMB maps, obtained from maximizing the angular momentum dispersion. The second last column gives the absolute value of the scalar-product between the orientation vectors of the quadrupole and the octopole. In an isotropic universe, the latter is uniformly distributed on the interval [0, 1]. The last column gives the probability of such an alignment (or stronger than that) to occur.

Method	(l,b) quadrupole [°]	(l,b) octopole [°]	ang. distance [°]	scalar-product	probability
C-R	(228.2,60.3)	(246.1,66.0)	9.80	0.985	0.019
NILC	(241.3,77.3)	(241.7,64.2)	13.1	0.974	0.033
SEVEM	(242.4,73.8)	(245.6,64.8)	9.08	0.988	0.016
SMICA	(238.5,76.6)	(239.0,64.3)	12.3	0.977	0.032

Cruz et al. (2011) found that a small region of the sky localised to both the ecliptic and Galactic south and near to the Galactic plane (their so-called gp10 region) exhibited particularly high variance. Thus, since the skewness is negative, we consider a prominent cold spot at ($b = -8°$, $l = 32°$), partially masked by the Galactic plane. However, when masking the seven coldest pixels of the spot, the significance of the skewness and kurtosis drops only slightly, with lower tail probabilities of approximately 0.03 and 0.93 respectively. If the whole gp10 region ($f_{\rm sky} = 7\%$) is masked, the skewness and kurtosis drop drastically and have lower tail probabilities of approximately 0.30 and 0.50 respectively, whereas the variance is highly significant since none of the 1000 simulations has a variance below the data. In order to check the possible leakage of Galactic contamination due to the Gaussian smoothing applied to the low resolution data, we repeated our calculations for the Wiener filtered maps used in Sect. 5.1, but found little variation to the existing results. Therefore, it is unlikely that any leakage impacts the estimators.

The incompatibility of the observed variance with simulations based on a cosmological model that has been determined from the same data set might appear puzzling at first, but can be understood as follows. The map-based variance is dominated by contributions from large angular scales on the sky, whilst the cosmological parameter fits are relatively insensitive to these low-order ℓ-modes, and are instead largely dominated by scales corresponding to $\ell > 50$. Thus, the best-fit spectrum in the context of a 6-parameter ΛCDM model can have a mismatch with the data on these scales, so that the corresponding simulations will not adequately capture the dearth of power at low-ℓ. The results presented here do indeed imply that the large-angular scale power is low relative to the fiducial sky model. In fact, when subtracting the quadrupole and octupole from both the data and simulations outside the U73 mask, the results are more consistent. In this case, the lower tail probabilities for the variance, skewness and kurtosis are 0.192, 0.637 and 0.792 respectively. This result was already found in Cruz et al. (2011). It is then plausible that the low multipole alignment could have the same cause as the anomalies considered here. However, when subtracting the quadrupole and octupole outside the CL58 mask, the lower tail probability for the low variance is 0.036, which remains rather low. The connection with the very low power in the ecliptic northern hemisphere also remains to be explored.

5.3. Hemispherical Asymmetry

In Eriksen et al. (2004a) and Hansen et al. (2004) it was discovered that the angular power spectrum of the first year *WMAP* data, when estimated locally at different positions on the sphere, appears not to be isotropic. In particular, the power spectrum calculated for a hemisphere centered at $(\theta, \phi) = (110°, 237°)$ (in Galactic co-latitude and longitude) was larger than when calculated in the opposite hemisphere over the multipole range $\ell = 2 - 40$. Simultaneously, Park (2004) also presented evidence for the existence of such hemispherical asymmetry — in which a particular statistical measure is considered to change discontinuously between two hemispheres on the sky — with the application of Minkowski functionals to the *WMAP* data. Since the preferred direction of Eriksen et al. (2004a) lies close to the ecliptic plane, it was also demonstrated that the large-angular scale N-point correlation functions showed a difference in behaviour when computed on ecliptic hemispheres. Many studies have subsequently been undertaken focusing on hemispheres in the ecliptic coordinate system, with Schwarz et al. (2004) particularly emphasizing the connection. Hemispherical asymmetry has also been seen with other measures of non-Gaussianity (Eriksen et al. 2004c, 2005; Räth et al. 2007a).

Table 18. Lower tail probablity for the variance, skewness and kurtosis at $N_{\rm side} = 16$, using different masks.

Mask	C-R	NILC	SEVEM	SMICA
		Variance		
U73, $f_{\rm sky} = 78\%$	0.019	0.017	0.014	0.019
CL58, $f_{\rm sky} = 58\%$	0.004	0.003	0.003	0.003
CL37, $f_{\rm sky} = 37\%$	0.028	0.017	0.018	0.016
Ecliptic North, $f_{\rm sky} = 39\%$	0.001	0.001	0.001	0.002
Ecliptic South, $f_{\rm sky} = 39\%$	0.464	0.479	0.454	0.490
		Skewness		
U73, $f_{\rm sky} = 78\%$	0.016	0.015	0.023	0.012
CL58, $f_{\rm sky} = 58\%$	0.208	0.139	0.162	0.147
CL37, $f_{\rm sky} = 37\%$	0.517	0.467	0.503	0.469
Ecliptic North, $f_{\rm sky} = 39\%$	0.502	0.526	0.526	0.521
Ecliptic South, $f_{\rm sky} = 39\%$	0.004	0.006	0.008	0.004
		Kurtosis		
U73, $f_{\rm sky} = 78\%$	0.972	0.973	0.966	0.982
CL58, $f_{\rm sky} = 58\%$	0.630	0.726	0.711	0.711
CL37, $f_{\rm sky} = 37\%$	0.069	0.135	0.130	0.124
Ecliptic North, $f_{\rm sky} = 39\%$	0.094	0.229	0.196	0.245
Ecliptic South, $f_{\rm sky} = 39\%$	0.933	0.916	0.886	0.948

Here we repeat the analysis of Eriksen et al. (2005) on the *Planck* component separated data, smoothed and then downgraded to $N_{\rm side} = 64$ as described in Sect. 2. As already noted in Sect. 4.3, the results for the low resolution maps are the most deviant relative to the MC simulations based on the *Planck* fiducial model.

The N-point correlation functions computed on the northern and southern hemispheres determined in the ecliptic coordinate frame and using the U73 mask are shown in Fig. 23. The correlation functions for the four *Planck* maps are very consistent, and the observed behaviour is in agreement with that seen in

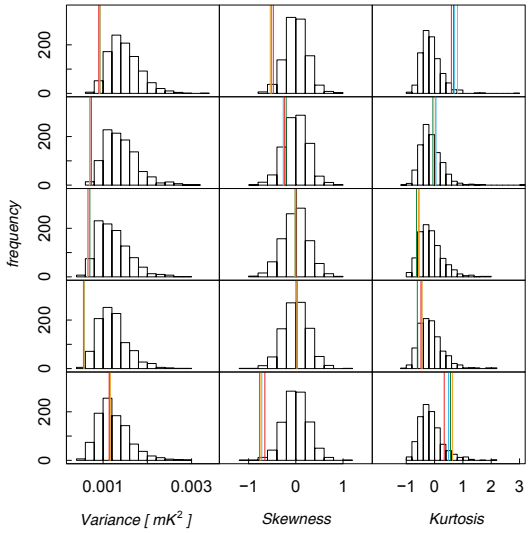


Fig. 21. Variance, skewness and kurtosis at $N_{\rm side} = 16$, for the U73 mask, CL58, CL37, ecliptic North and ecliptic South (from top to bottom). The different lines represent the four component separation methods C-R (green), NILC (blue), SEVEM (red), and SMICA (orange).

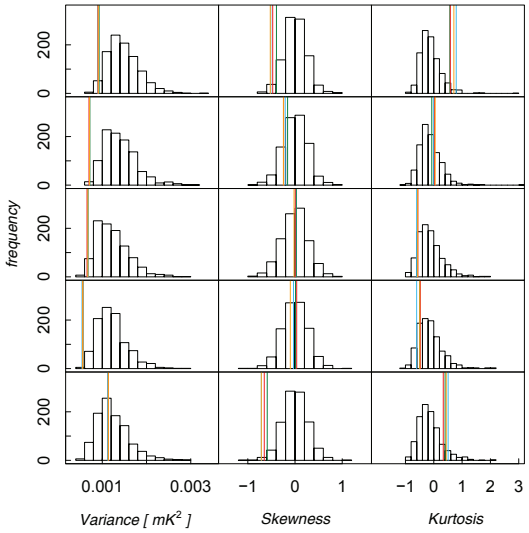


Fig. 22. Variance, skewness and kurtosis at $N_{\rm side} = 16$, for the U73 mask, CL58, CL37, ecliptic North and ecliptic South (from top to bottom). The different lines represent the four considered frequencies, namely 70 GHz (green), 100 GHz (blue), 143 GHz (red), and 217 GHz (orange).

Table 19. Frequency dependence of the lower tail probablity for the variance skewness and kurtosis at $N_{\rm side} = 16$, using different masks.

Mask	70 GHz	100 GHz	143 GHz	217 GHz
		Variance		
U73, $f_{\rm sky} = 78\%$	0.019	0.013	0.014	0.016
CL58, $f_{\rm sky} = 58\%$	0.003	0.003	0.003	0.003
CL37, $f_{\rm sky} = 37\%$	0.024	0.020	0.018	0.020
Ecliptic North, $f_{\rm sky} = 39\%$.	0.001	0.002	0.001	0.001
Ecliptic South, $f_{\rm sky} = 39\%$.	0.446	0.436	0.455	0.455
		Skewness		
U73, $f_{\rm sky} = 78\%$	0.045	0.016	0.024	0.015
CL58, $f_{\rm sky} = 58\%$	0.254	0.205	0.162	0.157
CL37, $f_{\rm sky} = 37\%$	0.503	0.471	0.468	0.515
Ecliptic North, $f_{\rm sky} = 39\%$.	0.505	0.447	0.541	0.352
Ecliptic South, $f_{\rm sky} = 39\%$.	0.015	0.006	0.009	0.006
		Kurtosis		
U73, $f_{\rm sky} = 78\%$	0.962	0.981	0.965	0.974
CL58, $f_{\rm sky} = 58\%$	0.619	0.684	0.710	0.725
CL37, $f_{\rm sky} = 37\%$	0.114	0.091	0.130	0.121
Ecliptic North, $f_{\rm sky} = 39\%$.	0.180	0.096	0.203	0.180
Ecliptic South, $f_{\rm sky} = 39\%$.	0.902	0.920	0.882	0.909

the *WMAP* data (Eriksen et al. 2004a). Specifically, the northern hemisphere correlation functions are featureless (both the three- and four-point functions lie very close to zero), whereas the southern hemisphere functions exhibit a level of structure that is in good agreement with the confidence regions computed from the Gaussian simulations.

The probabilities of obtaining a value for the χ^2 statistic for the *Planck* fiducial ΛCDM model at least as large as the observed values are presented in Table 20. The probabilities for the 3-point and 4-point functions in the northern Ecliptic hemisphere are especially large, and in the case of the pseudo-collapsed configuration all simulations yielded a larger than observed value of the χ^2. Nominally, this value is even more remarkable than found with the *WMAP* data (Eriksen et al. 2004a), although to interpret it correctly one has to keep in mind that the analysis presented here is an example of *a posteriori* statistic. Specific choices have been made about the smoothing scale used for downgrading the data, and, in particular, for the reference direction that defines the hemispheres. This will tend to overestimate the significance of the results. Nevetheless, the observed properties of the *Planck* data are consistent with a remarkable lack of power in a direction towards the north ecliptic pole, consistent with the simpler one-point statistics presented in Sect. 5.2.

5.4. Phase correlations

Previous studies using the methods of scaling indices and surrogates and based on the *WMAP* three-, five- and seven-year data (Räth et al. 2009, 2011; Rossmanith et al. 2012; Modest et al. 2013) showed significant evidence for intrinsic phase correlations at low ℓ values in the CMB. The signal was demonstarted to be robust with respect to the *WMAP* data release, to the component separation methods and to the selected test statistics. In this section we apply these methods to the *Planck* component separated data sets.

The scaling index method represents one way to estimate the local scaling properties of a point set in an arbitrary d-dimensional embedding space. The technique provides the possibility to reveal local structural characteristics of a given point distribution. A number of analyses have used scaling indices to test the Gaussian nature and statistical isotropy of the CMB

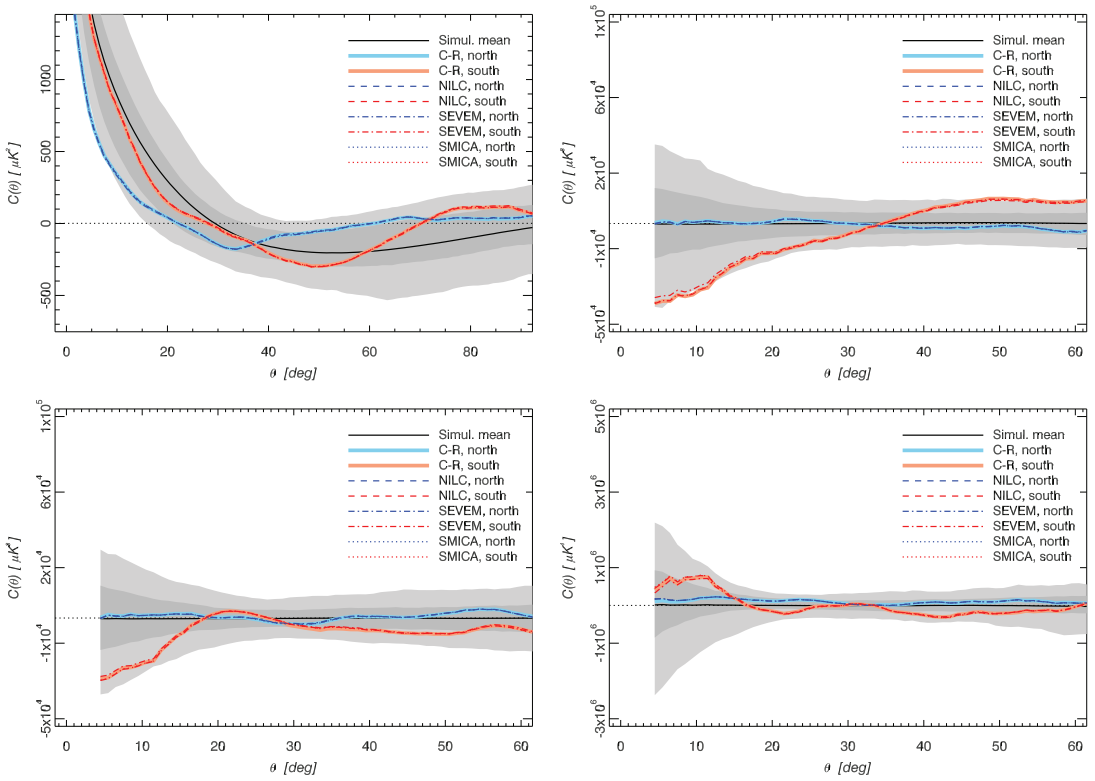


Fig. 23. The 2-point (upper left), pseudo-collapsed (upper right), equilateral 3-point (lower left) and rhombic 4-point (lower right) correlation functions ($N_{\rm side} = 64$). Correlation functions are shown for the analysis performed on northern (blue) and southern (red) hemispheres determined in the Ecliptic coordinate frame. The shaded dark and light grey bands indicate the 68% and 95% confidence regions, respectively.

as represented by the *WMAP* data (Räth et al. 2007a, 2009; Rossmanith et al. 2009a; Räth et al. 2011).

In general, the method is a mapping that calculates, for each member $p_i, i = 1, \ldots, N_{\rm pix}$ of a point set P, a single value that depends on the spatial position of $\boldsymbol{p}_i$ relative to the group of other points in its neighborhood, in which the point under consideration is embedded. A three-dimensional point set P is generated for two-dimensional spherical CMB-data by transforming the temperature values $T(\theta_i, \phi_i)$ of each pixel to a radial jitter around a sphere of radius R at the position of the pixel centre (θ_i, ϕ_i). For obtaining scaling indices the local weighted cumulative point distribution which is defined as

$$\rho(\boldsymbol{p}_i, r) = \sum_{j=1}^{N_{pix}} s_r(d(\boldsymbol{p}_i, \boldsymbol{p}_j)) \tag{34}$$

with r describing the scaling range, while s_r and d denote a shaping function and a distance measure, respectively, is calculated first. The scaling index $\alpha(\boldsymbol{p}_i, r)$ is then defined as the logarithmic derivative of $\rho(\boldsymbol{p}_i, r)$ with respect to r:

$$\alpha(\boldsymbol{p}_i, r) = \frac{\partial \log \rho(\boldsymbol{p}_i, r)}{\partial \log r} \,. \tag{35}$$

Using a quadratic gaussian shaping function $s_r(x) = e^{-(\frac{x}{r})^2}$ and an isotropic euclidian norm $d(\boldsymbol{p}_i, \boldsymbol{p}_j) = \|\boldsymbol{p}_i - \boldsymbol{p}_j\|$ as distance measure, one obtains the following analytic formula for the scaling indices

$$\alpha(\boldsymbol{p}_i, r) = \frac{\sum_{j=1}^{N_{rmpix}} 2\left(\frac{d_{ij}}{r}\right) e^{-\left(\frac{d_{ij}}{r}\right)^2}}{\sum_{j=1}^{N_{rmpix}} e^{-\left(\frac{d_{ij}}{r}\right)^2}} \,, \tag{36}$$

where we use the abbreviation $d_{ij} \equiv d(\boldsymbol{p}_i, \boldsymbol{p}_j)$. As should be clear from equation (36), the calculation of scaling indices depends on the scale parameter r. Ten scaling range parameters $r_k = 0.05, 0.1, \ldots, 0.5$, $k = 1, 2, \ldots 10$ in the notation of Räth et al. (2007a) are used in this analysis. In order to calculate scaling indices on large scales as in previous studies, we couple the r-jitter a to r_k via $a = 0.5 r_k$. The mean $\langle \alpha(r_k) \rangle$ and the standard deviation $\sigma_{\alpha(r_k)}$ derived from the full sky and from a set of 768 rotated hemispheres are used to test for non-Gaussianity and deviations from statistical isotropy.

In order to quantify the significance of the scaling index results, and focus the study on the phase properties of the observed CMB sky, we utilize the method of surrogate maps (Räth et al. 2009). Such a technique offers the unique possibility to

Table 20. Probabilities of obtaining values of the χ^2 statistic for the *Planck* fiducial model at least as large as the observed values of the statistic for the *Planck* maps with resolution parameter $N_{\rm side} = 64$ estimated using the C-R, NILC, SEVEM and SMICA methods.

	C-R	NILC	SEVEM	SMICA
Two-point function				
Northern Ecliptic	0.935	0.924	0.927	0.932
Southern Ecliptic	0.633	0.599	0.639	0.592
Pseudo-collapsed three-point function				
Northern Ecliptic	1.000	1.000	1.000	1.000
Southern Ecliptic	0.349	0.310	0.381	0.301
Equilateral three-point function				
Northern Ecliptic	0.996	0.999	0.999	0.999
Southern Ecliptic	0.627	0.644	0.678	0.656
Rhombic four-point function				
Northern Ecliptic	0.999	0.999	0.999	0.999
Southern Ecliptic	0.559	0.548	0.574	0.553

test for scale-dependent non-Gaussianity and deviations from isotropy in a completely model-independent ("blind") way. This self-consistency of the surrogate approach suppresses the sensitivity of the null tests to the assumed fiducial power spectrum. This is particularly pertinent given the potential mismatch of the *Planck* data to the fiducial spectrum on large-angular scales (Planck Collaboration XV 2013). The statistical properties of a Gaussian random field on the sphere can be fully described by its two-point correlation function (or power spectrum) and exhibit Fourier phases that are independent and identically distributed (i.i.d.) and follow a uniform distribution in the interval $[-\pi, \pi]$. Thus, demonstrating the existence of Fourier phase correlations in CMB maps could indicate the presence of non-Gaussianities in the primordial density fluctuations. The possible presence of phase correlations is tested using the method of surrogates.

However, the Gaussianity of the temperature distribution and the randomness of the set of Fourier phases of the map to be studied are a necessary pre-requisite for the application of the surrogate-generating algorithm. Therefore the following preprocessing steps are applied to generate a zero order surrogate map. First, the maps are remapped on to a Gaussian distribution in a rank-ordered way. Then we ensure the randomness of the set of Fourier phases by making a rank-ordered remapping of the phases on to a set of uniformly distributed ones. These two preprocessing steps only have marginal influence on the maps. Now, the set of surrogates to be used for assessing the statistical properties of the data sets can be generated by shuffling the phases in the space of the spherical harmonics while exactly preserving the modulus of the $a_{\ell m}$. Moreover, by introducing a two-step shuffling scheme for previously specified ℓ-ranges, a scale-dependent analysis is made possible. It is worth noticing that while in all surrogate maps the modulus of the $a_{\ell m}$ is exactly preserved, null tests involving a comparison to an assumed fiducial power spectrum only preserve the C_ℓ values, which are average values of the $|a_{\ell m}|$ when summed over m. Thus, the linear properties of the surrogate maps are more tightly constrained, and specifically kept constant, than in tests involving simulated maps generated on the basis of the C_ℓs.

So-called first and second order surrogate maps are then obtained as follows. We initally generate a first order surrogate map, in which any phase correlations for the scales that are not of interest are randomized by shuffling its phases $\phi_{\ell m}$ for $\ell \notin \Delta l = [\ell_{\rm min}, \ell_{\rm max}], 0 < m \leq \ell$. In a second step, N ($N = 1000$ throughout these investigations) realizations of second order surrogate maps are generated from the first order surrogate map, in which the remaining phases $\phi_{\ell m}$ with $\ell \in \Delta\ell$, $0 < m \leq \ell$ are shuffled, while the previously randomized phases for the other scales are preserved. The generation of surrogates is always performed using the maps with the highest resolution, i.e., $N_{\rm side} = 2048$. Given the evidence for anomalies on the largest angular scales, and to ensure consistency with the previous *WMAP* analyses, we perform dedicated scale-dependent tests for the scales defined by $\Delta\ell = [2, 20]$.

Since the methodology in its simplest form requires the orthonormality of the set of basis functions $Y_{\ell m}$, we apply the method to the full sky foreground-cleaned maps as obtained after component separation with Commander-Ruler, NILC, SEVEM and SMICA. For the selected ℓ-interval $\Delta\ell = [2, 20]$, the generation of the first order surrogate map removes the phase signature of the small scale residuals in the data and can be interpreted as a form of inpainting procedure for small masked patches in the Galactic plane. The differences between the first and second order surrogates are quantified by the σ-normalized deviation S

$$S(Y) = \frac{Y_{\rm surro1} - \langle Y_{\rm surro2}\rangle}{\sigma_{Y_{\rm surro2}}} \tag{37}$$

with, $Y = \langle\alpha(r_k)\rangle, \sigma_{\alpha(r_k)}, \chi^2$. Here, χ^2 represents either a diagonal combination of the the mean $\langle\alpha(r_k)\rangle$ and standard deviation $\sigma_{\alpha(r_k)}$ at a certain scale r_k or for the full scale-independent χ^2 statistics

$$\chi^2 = (\boldsymbol{M} - \langle\boldsymbol{M}\rangle)^{\rm T}\mathsf{C}^{-1}(\boldsymbol{M} - \langle\boldsymbol{M}\rangle), \tag{38}$$

where the test statistics to be combined are comprised in the vector $\boldsymbol{M}$ and C is obtained by cross correlating the elements of $\boldsymbol{M}$. With the mean and the standard deviation as input for $\boldsymbol{M}$ we obtain $\chi^2_{\langle\alpha\rangle}$ and $\chi^2_{\sigma_\alpha}$ statistics with $\boldsymbol{M}^{\rm T} = (\langle\alpha(r_1)\rangle, \ldots, \langle\alpha(r_{10})\rangle)$ and $\boldsymbol{M}^{\rm T} = (\sigma_{\alpha(r_1)}, \ldots, \sigma_{\alpha(r_{10})})$ respectively.

Fig. 24 shows the $S(\chi^2)$ values for the set of rotated hemispheres for the SMICA map. Each pixel of the full sky map with a HEALPix resolution of $N_{\rm side} = 8$ specifies one of the 768 S-values for a rotated hemisphere, where the pixel position indicates the orientation of the z-axis of the rotated coordinate system. We detect pronounced signatures for both non-Gaussianities and anisotropies. The results are consistent for Commander-Ruler, NILC, SEVEM and SMICA.

In Fig. 25, the deviations $S(Y)$ are displayed for the mean and standard deviation. We only show the results for the SMICA map. The other three maps yield very similar results. For all four maps the values for $S(\langle\alpha\rangle)$ extend far beyond 3 for $r = 0.2 - -0.25$ when rotated hemispheres are considered separately. Since the effect in the separate hemispheres goes in opposing directions, no signal is observed for the full sky. The results for the scale-independent χ^2 statistics are summarized in Table 21. The results suggest a highly significant detection of both non-Gaussianities and anisotropies in the *Planck* data, consistent with those obtained previously with *WMAP* data (for comparison see Modest et al. 2013).

We have also investigated whether the significance of the results depends on the choice of $\ell_{\rm min}$ and $\ell_{\rm max}$. In particular, we have extended the range of interest to $\ell_{\rm max} = 30$, and then considered three sub-intervals, $\Delta\ell = [2, 10]$, $\Delta\ell = [11, 20]$ and

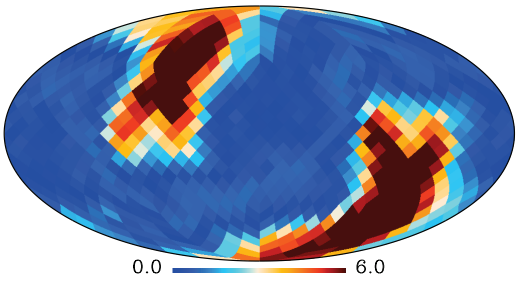


Fig. 24. Deviations $S(\chi^2)$ of the rotated hemispheres derived from a combination of the mean and the standard deviation of the scaling indices for the scale r_5 determined from the SMICA map.

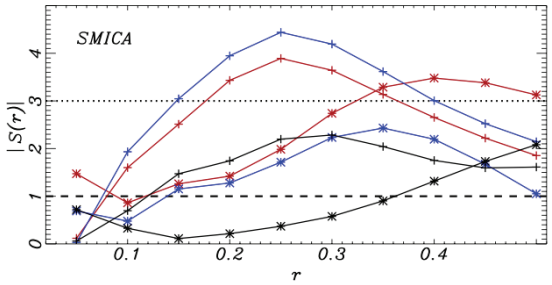


Fig. 25. Deviations $|S(r)|$ for the SMICA map as a function of the scale parameter r for the full sky (black) and upper (red) and lower (blue) rotated hemispheres. The plus signs denote the results for the mean $\langle\alpha(r_k)\rangle$, the star-signs represent the standard deviation $\sigma_{\alpha(r_k)}$. The dashed (dotted) line indicates the 1 (3) σ significance interval.

Table 21. Deviations S and empirical probabilities p for the scale-independent χ^2-statistics derived from the C-R, NILC, SEVEM and SMICA maps.

	Full Sky ($S/\%$)	Upper Hemisphere ($S/\%$)	Lower Hemisphere ($S/\%$)
C-R, $\chi^2_{\langle\alpha\rangle}$	0.86 / 82.6	4.21 / 99.7	3.18 / 99.0
C-R, $\chi^2_{\sigma_\alpha}$	0.88 / 85.2	3.94 / 99.5	3.10 / 99.2
NILC, $\chi^2_{\langle\alpha\rangle}$	0.86 / 81.8	3.74 / 99.6	4.41 / >99.9
NILC, $\chi^2_{\sigma_\alpha}$	0.79 / 78.8	3.69 / 99.6	4.49 / >99.9
SEVEM, $\chi^2_{\langle\alpha\rangle}$	0.00 / 58.0.	3.22 / 99.3	5.02 / >99.9
SEVEM, $\chi^2_{\sigma_\alpha}$	0.05 / 60.8	3.20 / 99.0	5.11 / 99.9
SMICA, $\chi^2_{\langle\alpha\rangle}$	0.75 / 80.1	3.80 / 99.8	4.70 / 99.8
SMICA, $\chi^2_{\sigma_\alpha}$:	0.01 / 54.4	3.64 / 99.3	4.81 / >99.9

$\Delta\ell = [21, 30]$. Over the full range, similar results are found to those from $\Delta\ell = [2, 20]$, but at lower significance. This suggests that the inclusion of the phases in the interval $\Delta\ell = [21, 30]$ dilutes the signal, because there are no phase correlations in this ℓ-range. This is corroborated by the fact that the first and second order surrogates generated specifically for this sub-interval cannot be distinguished. The results for the interval $\Delta\ell = [2, 10]$ are quite consistent with those over $\Delta\ell = [2, 20]$, whereas for $\Delta\ell = [11, 20]$ we find that the signature in the northern ecliptic hemisphere nearly vanishes. Conversely, in the southern ecliptic hemisphere, on the other hand, the S-signal persists – especially for regions covering the Cold Spot. It thus appears that the lowest ℓ-range is predominantly responsible for the detected hemispherical asymmetries detected in the spectrum of scaling indices, whilst the interemediate interval considered may have an association with the Cold Spot. It is certainly the case that scale-dependence is seen in the nature of the phase correlations present in the data.

Since both, the modulus of the $a_{\ell m}$s for all ℓs and the phases $\phi_{\ell m}$ for $\ell \notin \Delta\ell$ are exactly the same in the first and second order surrogates, one must infer that the pattern of hemispherical asymmetry in the S-maps can solely be attributed to phase correlations in the interval $\Delta\ell$. Thus, the analysis involving surrogate maps reveals that there *are* phase correlations at low ℓ.

5.5. Dipolar asymmetry

In previous sections, we have seen evidence for a break in isotropy related to the discontinuous distribution of power in hemispheres on the sky. Bennett et al. (2011) distinguishes between such an asymmetry and one where the CMB signal is modulated across the sky by a dipolar term. Studies of such a dipolar asymmetry have been motivated by the phenomenological proposal of Gordon et al. (2005) that the power asymmetry could be described in terms of a multiplicative dipole modulation model. In addition, relativistic Doppler boosting due to our motion with respect to the CMB rest frame is expected to induce a dipolar modulation aligned with the CMB dipole at the $O(10^{-3})$ level; a statistically significant detection of this effect by *Planck* is presented by Planck Collaboration XXVII (2013).

5.5.1. Power asymmetry

In their analysis of the 5-year *WMAP* data, (Hansen et al. 2009) specifically searched for the dipolar modulation of power on the sky, In particular, a simple test was performed in which the power spectrum on discs was computed and binned into independent blocks of 100 multipoles from $\ell = 2$ to $\ell = 600$, then each block fitted for a dipolar asymmetry in the power distribution. The 6 ℓ-ranges considered showed evidence of a consistent dipole direction, yet, from a set of 10000 simulations, none showed a similarly strong asymmetry. A further extension of the analysis introduced a model selection procedure taking into account the statistical penalty for introducing an asymmetric model with additional parameters (direction of asymmetry, amplitude of asymmetry and asymmetric multipole ranges). Even in this case, the asymmetry was found to be highly significant for the whole range $\ell = 2 - 600$.

However, such a procedure is highly expensive in terms of CPU-time. Given the higher sensitivity and angular resolution of the *Planck* data, we have therefore elected to focus on the simpler disc-based test, thus allowing us to probe further into a previously unexplored ℓ-range. This should at least in part answer any *a posteriori* criticisms of the study. Since the analysis is power-spectrum based, the half-ring data sets for the different CMB estimators are used. The approach is as follows:

1. The half-ring temperature maps are multiplied with an appropriate Galactic and point source mask.

2. The cross power spectrum between the two halves is estimated locally using the MASTER approach (Hivon et al. 2002) for 768 discs of diameter 22.5° degrees centered on the pixel centers of all the pixels of an $N_{\rm side} = 8$ HEALPix map.
3. We apply the same procedure to the set of 500 simulated maps of CMB and noise.
4. The 768 local spectra are binned into blocks of about 100 multipoles as in Hansen et al. (2009). There are not exactly 100 multipoles in each block, as the spectra have been initially estimated in 16-multipole blocks.
5. For each 100-multipole block and each disc, the mean power from simulations is subtracted and the result is divided by the standard deviation. Dividing the spectra by the local standard deviation avoids the problem that directions close to the Galaxy, where the Galactic mask increases the variance, can dominate the statistics due to large fluctuations.
6. Each 100-multipole block now has an associated map at $N_{\rm side} = 8$, where each pixel corresponds to the normalised power spectrum estimated on the disc centered on that direction.
7. Spherical harmonic transforms are computed for each of the maps in order to obtain the dipoles and the dipole directions (θ, ϕ) of the power asymmetry for each 100-multipole block. The alignment of this direction between the different multipole blocks is then a measure of the power spectrum asymmetry. Despite the mask-induced correlations between adjacent multipoles, the power spectra estimated in 100-multipole blocks should be independent and the dipole directions of an isotropic field should be random.

In order to assess the significance of the asymmetry, one has to find out whether the distribution of dipole directions for the different scales are as random and independent as in the simulated maps. For this purpose, we define a dispersion angle, $\theta_{\rm mean}$, which is the mean angle between all possible combinations of 100-multipole dipole directions up to a given $\ell_{\rm max}$. We calculate $\theta_{\rm mean}(\ell_{\rm max} = 1500)$ for the data and compare it to the simulations.

Table 22 presents a summary of the power asymmetry results from the *Planck* data processed by the four foreground cleaning methods — Commander-Ruler, NILC, SEVEM and SMICA — computed on the U73 mask. To illustrate the effect of the mask in the analysis, we also show the result obtained using the SMICA data with a smaller mask with $f_{\rm sky} = 88\%$ (the CS-SMICA89 mask which corresponds to the confidence mask for that method – see Planck Collaboration XII 2013). For comparison, we have also included the latest *WMAP* 9-year result computed with their KQ85 mask.

From Table 22 we see that the result from the mask with larger sky coverage, the SMICA data with the CS-SMICA89 mask, has the highest power asymmetry, the data dispersion angle of the first 15 100-multipole dipole directions being lower than all the 500 simulations. The significance decreases to about 99.2% confidence level, however, for the U73 mask with a smaller sky coverage, except for the case of Commander-Ruler, which has an even lower significance. Moreover, the dispersion angles among the first 15 100-multipole dipole directions for the four methods are consistent. The comparison between the simulations and the data dispersion angles up to $\ell_{\rm max} = 1500$ is shown in Fig. 26.

In Fig. 27 we show the dipole directions of the 15 initial 100-multipole bins for the SMICA map with the CS-SMICA89 mask, as well as the 6 first 100-multipole bins for *WMAP*9 data with the KQ85 mask (squares). We see that the direction of the first 6 dipoles are similar to the directions found in the *WMAP* data. The preferred directions for *WMAP*9 and *Planck* over the range $\ell = 2 - 600$ are indicated, together with the *Planck* direction for the total range $\ell = 2 - 1500$. Finally, the direction of the dipole modulation described in Sect. 5.5.2 is also included. Similar behaviour is seen for all of the *Planck* foreground cleaned maps and for the U73 mask, although the scatter between the dipole directions increases with increasing sky cut.

It should be apparent that the asymmetry direction from the largest to the smallest angular scales are in general tightly clustered around the same direction as found for *WMAP*. However, with the *Planck* data a second preferred direction is also seen which is aligned with the CMB dipole direction. This result is consistent with the findings of Planck Collaboration XXVII (2013), who reports a statistically significant detection of Doppler boosting aligned with the CMB dipole at small angular scales.

Table 22. Summary of the power dipole directions on the sky, up to $\ell_{\rm max} = 1500$, as determined from maps of the power spectrum estimated from 768 22.5° radius discs and averaged over $\Delta\ell = 100$ bins. The significance of the power asymmetry, shown in the last column, is quantified by the fraction of simulations that have smaller clustering of the dipole directions than the data. For the *Planck* analysis we used the 500 FFP6 simulations, while for *WMAP* we used 10000 Gaussian simulations.

Method	Mask	(l,b) [°]	$\theta^d_{\rm mean}$ [°]	Frac. $\theta^{\rm sim}_{\rm mean} < \theta^d_{\rm mean}$
C-R	U73	(231,−5)	67.8	11/500
NILC	U73	(223,−1)	66.1	4/500
SEVEM	U73	(224,2)	66.6	4/500
SMICA	U73	(225,1)	66.2	4/500
*WMAP*9	*WMAP*9 KQ85	(226,-27)	33.2	27/10000
SMICA	CS-SMICA89	(224,0)	55.8	0/500

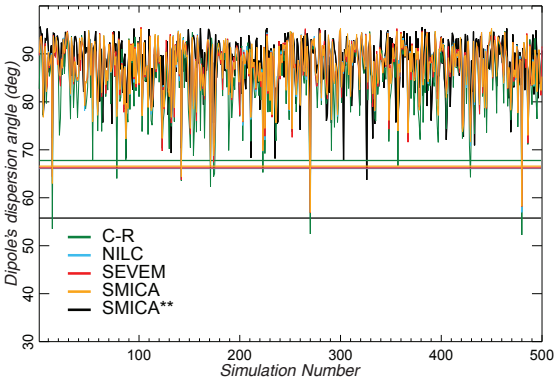


Fig. 26. Dispersion angles of the power spectra dipole directions, the mean of the differences of the dipole direction angles, up to $\ell_{\rm max} = 1500$, of the 500 FFP6 simulations compared to the *Planck* data with different foreground cleaning methods. All analyses, except SMICA**, are performed with the U73 mask. The SMICA** case is for SMICA data with the CS-SMICA89 mask.

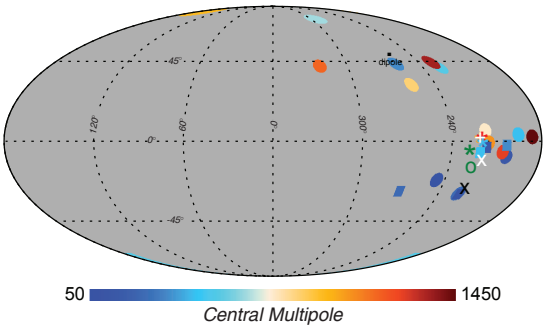


Fig. 27. Dipole directions for 100-multipole bins of the local power spectrum distribution from $\ell = 2 - 1500$ in the SMICA map with the CS-SMICA89 mask applied. We also show the total direction for $\ell_{\rm max} = 600$ for *WMAP*9 (black X) and SMICA (white X) as well as for $\ell_{\rm max} = 1500$ for SMICA (white big +). The stars with different colors correspond to C-R (green), NILC (deepskyblue), SEVEM (red) and SMICA (orange) with the U73 mask. The best fit dipole modulation direction from Sect. 5.5.2 is indicated by the white open circle.

In Fig. 28 we show the C_ℓ computed in discs of diameter 90° centered on the preferred asymmetry dipole direction for $\ell = 2-1500$ as well as the opposite direction. We can clearly see that one spectrum lies systematically above the other over the full multipole range, but in particular for the lowest multipoles. Such an asymmetry is not seen at the same level of significance when the spectra are computed for discs centred on the cosmological dipole direction.

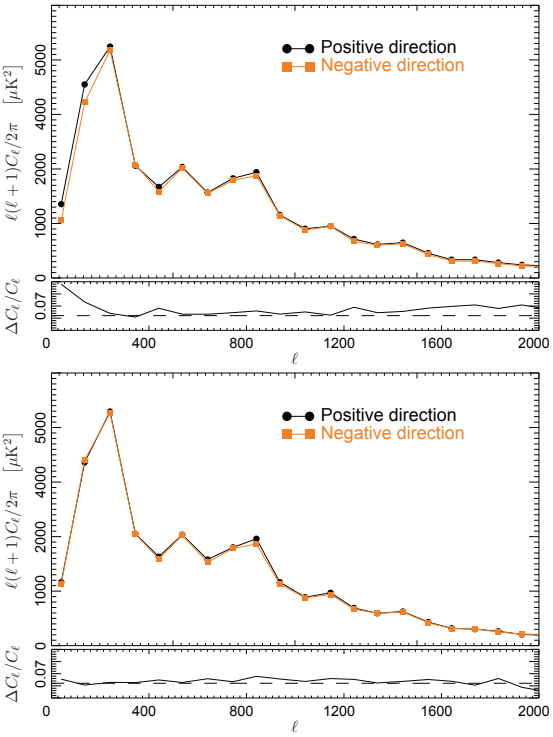


Fig. 28. *Top:* The power spectra calculated on discs with diameter 90° for the range $\ell = 2 - 1500$ in the direction of maximal asymmetry and its opposite. *Bottom:* The equivalent plot for the direction defined by the comological dipole. The lower panels indicate the normaised difference of the spectra from opposing directions.

5.5.2. Dipole modulation

In Sect. 5.5.1 it was shown that the previously reported power asymmetry is visible at all multipoles probed by *Planck* with a fairly consistent preferred axis across angular scales. No explicit parametric model was assumed in the analysis. In this section, however, we revisit the phenomenological model due to Gordon et al. (2005) considering only large angular scales, who proposed that the power asymmetry could be described in terms of a multiplicative dipole modulation model of the form $\mathbf{d} = (1 + A\,\boldsymbol{p}\cdot\boldsymbol{n})\mathbf{s}_{\rm iso} + \mathbf{n} \equiv \mathbf{M}\mathbf{s}_{\rm iso} + \mathbf{n}$, where A is the dipole amplitude, $\boldsymbol{p}$ is the dipole direction, $\mathbf{n}$ denotes instrumental noise, and $\mathbf{s}_{\rm iso}$ is an underlying isotropic CMB field. Both $\mathbf{s}_{\rm iso}$ and $\mathbf{n}$ are assumed to be Gaussian random fields with covariance matrices S and N, respectively. Since $\mathbf{s}_{\rm iso}$ is assumed to be isotropic, its covariance may be fully specified by some angular power spectrum $C_{\ell,{\rm iso}}$.

In the following we present the results from a direct likelihood analysis of this model, similar to those described by Eriksen et al. (2007a); Hoftuft et al. (2009) for the 3- and 5-year *WMAP* data, respectively. Since this method requires matrix inversions and determinant evaluations, the computational expense scales as $O(N_{\rm pix})$, and it is therefore feasible only at low resolutions. Specifically, we consider maps downgraded to a HEALPix pixel resolution of $N_{\rm pix} = 32$, smoothed to angular resolutions ranging from 5 to 10°, ensuring sufficient bandwidth limitation at this pixelization. All four *Planck* CMB solutions are included in the analysis; however, note that the Galactic plane is handled differently in the four approaches. Specifically, for the Commander map the region inside the corresponding analysis mask has been replaced with a Gaussian constrained realization, eliminating the possibility of bright Galactic residuals to leak outside the mask during degradation (Planck Collaboration XV 2013); for SMICA and NILC a smaller region is replaced with a Wiener filter; while for SEVEM no special precautions are taken.

After degrading each map to the appropriate resolution, we add random uniform Gaussian noise of 1μK rms to each pixel to regularize the covariance matrix. All pixels inside the U73 mask are excluded, and we adopt the difference maps between the raw *Planck* LFI 30 GHz and HFI 353 GHz maps and the SMICA CMB solution as two foreground templates, tracing low- and high-frequency foregrounds, respecively. We marginalize over these Galactic foreground templates, $\boldsymbol{f}$, as well as four monopole and dipole templates, by adding corresonding term of the form $\alpha\boldsymbol{f}\boldsymbol{f}^{\rm T}$ to the total data covariance matrix, where α is set to a numerically large value.

Before writing down the likelihood for A and $\boldsymbol{p}$, a choice has to made for the power spectrum, $C_{\ell,{\rm iso}}$. We follow Eriksen et al. (2007a), and adopt a simple two-parameter amplitude-tilt parameter model on the form $C_{\ell,{\rm iso}} = q\left(\ell/\ell_{\rm pivot}\right)^n C_{\ell,{\rm fid}}$ for this purpose, where the fiducial spectrum, $C_{\ell,{\rm fid}}$, is the best-fit *Planck* spectrum. The full model therefore includes five free parameters,

Table 23. Summary of dipole modulation likelihood results as a function of scale for all four *Planck* CMB solutions.

Data set	FWHM [°]	A	(l,b) [°]	$\Delta \ln \mathcal{L}$	Significance
Commander	5	$0.078^{+0.020}_{-0.021}$	$(227, -15) \pm 19$	8.8	3.5σ
NILC	5	$0.069^{+0.020}_{-0.021}$	$(226, -16) \pm 22$	7.1	3.0σ
SEVEM	5	$0.066^{+0.021}_{-0.021}$	$(227, -16) \pm 24$	6.7	2.9σ
SMICA	5	$0.065^{+0.021}_{-0.021}$	$(226, -17) \pm 24$	6.6	2.9σ
*WMAP*5 ILC	4.5	0.072 ± 0.022	$(224, -22) \pm 24$	7.3	3.3σ
Commander	6	$0.076^{+0.024}_{-0.025}$	$(223, -16) \pm 25$	6.4	2.8σ
NILC	6	$0.062^{+0.025}_{-0.026}$	$(223, -19) \pm 38$	4.7	2.3σ
SEVEM	6	$0.060^{+0.025}_{-0.026}$	$(225, -19) \pm 40$	4.6	2.2σ
SMICA	6	$0.058^{+0.025}_{-0.027}$	$(223, -21) \pm 43$	4.2	2.1σ
Commander	7	$0.062^{+0.028}_{-0.030}$	$(223, -8) \pm 45$	4.0	2.0σ
NILC	7	$0.055^{+0.029}_{-0.030}$	$(225, -10) \pm 53$	3.4	1.7σ
SEVEM	7	$0.055^{+0.029}_{-0.030}$	$(226, -10) \pm 54$	3.3	1.7σ
SMICA	7	$0.048^{+0.029}_{-0.029}$	$(226, -11) \pm 58$	2.8	1.5σ
Commander	8	$0.043^{+0.032}_{-0.029}$	$(218, -15) \pm 62$	2.1	1.2σ
NILC	8	$0.049^{+0.032}_{-0.031}$	$(223, -16) \pm 59$	2.5	1.4σ
SEVEM	8	$0.050^{+0.032}_{-0.031}$	$(223, -15) \pm 60$	2.5	1.4σ
SMICA	8	$0.041^{+0.032}_{-0.029}$	$(225, -16) \pm 63$	2.0	1.1σ
Commander	9	$0.068^{+0.035}_{-0.037}$	$(210, -24) \pm 52$	3.3	1.7σ
NILC	9	$0.076^{+0.035}_{-0.037}$	$(216, -25) \pm 45$	3.9	1.9σ
SEVEM	9	$0.078^{+0.035}_{-0.037}$	$(215, -24) \pm 43$	4.0	2.0σ
SMICA	9	$0.070^{+0.035}_{-0.037}$	$(216, -25) \pm 50$	3.4	1.8σ
*WMAP*3 ILC	9	0.114	$(225, -27)$	6.1	2.8σ
Commander	10	$0.092^{+0.037}_{-0.040}$	$(215, -29) \pm 38$	4.5	2.2σ
NILC	10	$0.098^{+0.037}_{-0.039}$	$(217, -29) \pm 33$	5.0	2.3σ
SEVEM	10	$0.103^{+0.037}_{-0.039}$	$(217, -28) \pm 30$	5.4	2.5σ
SMICA	10	$0.094^{+0.037}_{-0.040}$	$(218, -29) \pm 37$	4.6	2.2σ

namely three dipole parameters and two power spectrum parameters.

Taking advantage of the fact that both the signal and noise are assumed Gaussian, the exact likelhiood may be written down in a convenient closed form,

$$\mathcal{L}(A, \boldsymbol{p}, q, n) \propto \frac{e^{-\frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\left(\mathsf{M}^{\mathrm{T}}\mathsf{S}\mathsf{M}+\mathsf{N}+\alpha\sum_i f_i f_i^{\mathrm{T}}\right)^{-1}\boldsymbol{d}}}{\sqrt{|\mathsf{M}^{\mathrm{T}}\mathsf{S}\mathsf{M}+\mathsf{N}+\alpha\sum_i \boldsymbol{f}_i \boldsymbol{f}_i^{\mathrm{T}}|}}. \quad (39)$$

This expression forms the basis of all calculations presented in the following.

Due to the high computational expense associated with these evaluations, we do not compute the full joint five-parameter model in this analysis, only conditionals of it. However, we iterate once in a Gibbs-sampling like approach, by maximizing each conditional to obtain an approximation to the full maximum-likelihood solution. That is, we first map out the dipole likelihood for the 5° FWHM case, fixing the power spectrum at the fiducial spectrum, $\mathcal{L}(A, \boldsymbol{p}|q = 1, n = 0)$, and locate the maximum-likelihood dipole parameters. Then we map out the corresponding power spectrum conditional, $\mathcal{L}(q, n|A_{m\ell}, \boldsymbol{p}_{m\ell})$. Finally, we update the dipole likelihood with these power spectrum parameters, and evaluate the final results. Note that the power spectrum and dipole modulation parameters are only weakly correlated, and this procedure is therefore close to optimal. Further, the approach is also conservative, in the sense that it will always underestimate the significance of the dipole modulation model; the derived maximum-likelihood value will always lie slightly below the true maximum-likelihood point.

The results from these calculations are summarized in Table 23, listing results for all four *Planck* CMB maps at angular scales between 5 and 10° FWHM. For easy reference, we also list the results from the corrsponding 3- and 5-year *WMAP* analyses (Eriksen et al. 2007a; Hoftuft et al. 2009). Note that the former was performed at a HEALPix resolution of $N_{\mathrm{side}} = 16$ and the latter at an angular resolution of 4.5° FWHM.

Fig. 30 shows marginals for A, q and n, as derived from the Commander CMB solution for all smoothing scales. At least two interesting points can be seen here. First, while there is clearly significant scatter in the derived dipole modulation amplitude for different smoothing scales, as originally pointed out by Hanson & Lewis (2009), all curves appear to be consistent with a single value of $A \sim 0.07$. No other single value fits all scales equally well. Second, it is interesting to note that the low-ℓ power spectrum derived here is consistent, but not without some tension, with the fiducial spectrum, $(q, n) = (1, 0)$, around $1.5 - 2\sigma$. In

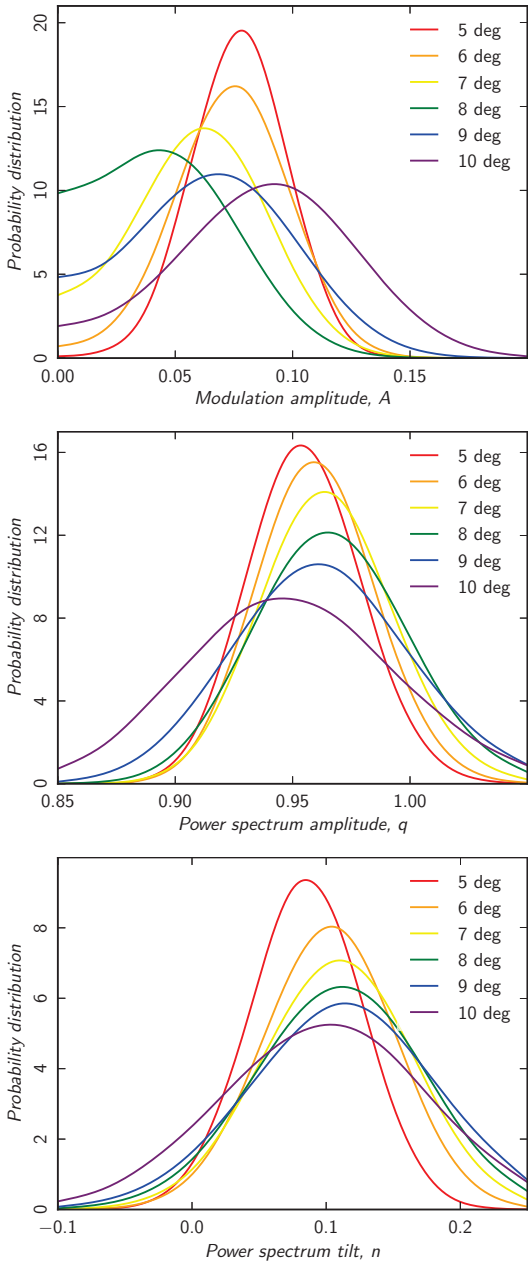


Fig. 29. Marginal dipole modulation amplitude (top), power spectrum amplitude (middle) and power spectrum tilt (bottom) probability distributions as a function of smoothing scale, shown for the `Commander` CMB solution.

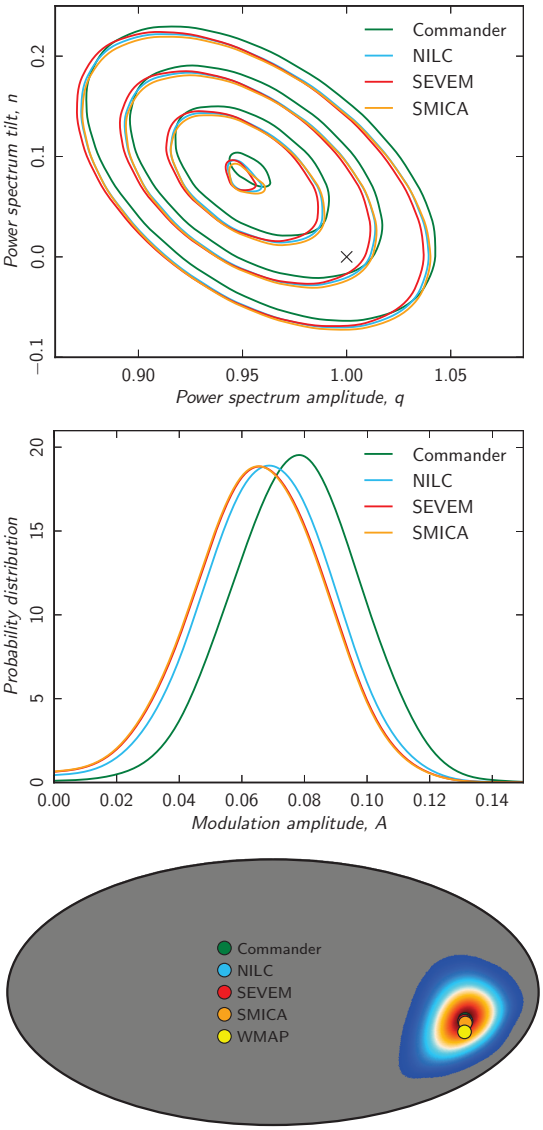


Fig. 30. Consistency between component separation algorithms as measured by the dipole modulation likelihood. The top panel shows the marginal power spectrum amplitude for the 5° smoothing scale, the middle panel shows dipole modulation amplitude, and the bottom panel shows the preferred dipole directions. The coloured area indicates the 95% confidence region for the `Commander` solution, while the dots shows the maximum-posterior directions for the other codes.

particular, there appears to be a slight trend toward a steeper and positive spectral index as more weight is put on the larger scales, a result already noted by *COBE*-DMR. The same conclusion is reached using the low-ℓ *Planck* likelihood, as described in Planck Collaboration XV (2013).

In Fig. 30 we compare the results from all four CMB solutions for the 5° FWHM smoothing scale. Clearly the results are consistent despite the use of different algorithms and different treatments of the Galactic plane, demonstrating robustness with

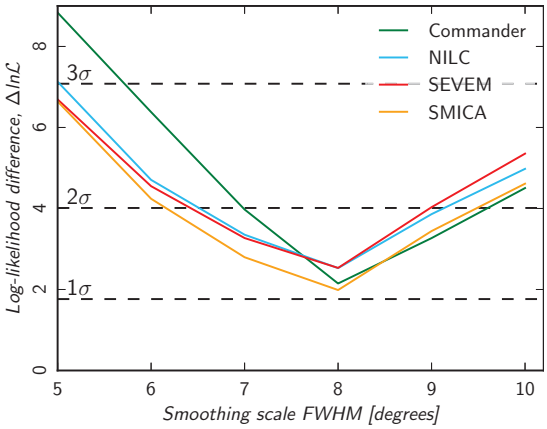


Fig. 31. Log-likelihood difference between the best-fit dipole modulation model and the fiducial isotropic model as a function of smoothing scale. Horizontal dashed lines indicate 1, 2 and 3σ thresholds.

respect to the details of the analysis methods. Further, we also note that these results are consistent with those derived from the 5-year *WMAP* ILC map by Eriksen et al. (2007a), demonstrating robustness across experiments. On the other hand, it is notable that a higher dipole amplitude was found at 9° FWHM for the 3-year *WMAP* ILC map than is observed here, using a larger mask.

In Fig. 31 we show the log-likelihood difference between the derived maximum-likelihood point and the isotropic model, $A = 0$, as a function of smoothing scale. The power spectrum parameters are kept fixed at the best-fit values for both points, leaving three additional parameters for the dipole model. The dashed horizontal lines indicate the 1, 2 and 3σ confidence regions for three degrees of freedom. As has been noted previously in the literature, these significances vary with smoothing scale. Taken at face value, the results presented here are suggestive but clearly not decisive, resulting in an unchanged situatio with respect to earlier reports. This is of course not unexpected given that *WMAP* is already strongly cosmic variance limited at these angular scales.

The critical question is whether the trend seen at smaller angular scales in Fig. 31 continues, or if the apparent likelihood peak at 5° FWHM happens to be a local maximum. Hanson & Lewis (2009), and later Bennett et al. (2011), address this question through a computationally cheaper quadratic estimator, allowing them to extend a similar analysis to small scales. In doing so, they claim that the apparent likelihood peak at 5° is indeed a local maximum, and the evidence for the modulation model falls off when more data are included. In this respect, it should be noted that the dipole modulation model was originally proposed by Gordon et al. (2005) as a simple phenomenological characterization of the more general power asymmetry. In particular, it assumes that the modulation amplitude, A, is equally strong on all scales. From both the results shown in Sect. 5.5.1 and presented by Hanson & Lewis (2009); Bennett et al. (2011), this appears not to hold, as the fractional hemispherical power difference is clearly smaller at $\ell > 300$ than at $\ell < 100$. On the other hand, the preferred directions derived from the current low-ℓ analysis is remarkably consistent with the high-ℓ direction derived in Sect. 5.5.1. A proper modulation model may therefore need additional spatial structure beyond the simple dipole proposed by Gordon et al. (2005), as already suggested by Hoftuft et al. (2009) and Moss et al. (2011).

5.6. Generalized modulation

In this section, we study a generalization of the dipolar modulation field analysed in section 5.5.2 using the Bipolar Spherical Harmonic (BipoSH) formalism. For a statistically isotropy sky, the spherical harmonic space two-point correlation matrix is diagonal, and, given by the angular power spectrum C_ℓ. The BipoSH representation provides a natural, mathematically complete, generalization of the angular power spectrum that captures statistical isotropy violations via coefficients that are a completely equivalent representation of the spherical harmonic correlation matrix,

$$A^{LM}_{\ell_1\ell_2} = \sum_{m_1 m_2} \langle a_{\ell_1 m_1} a_{\ell_2 m_2} \rangle C^{LM}_{\ell_1 m_1 \ell_2 m_2} \,. \tag{40}$$

This relationship combines the off-diagonal spherical harmonic correlations into a bipolar multipole L, M – analogous to the total angular momentum addition of states. The CMB angular power spectrum corresponds to the $L = 0$ BipoSH coefficients $C_\ell = (-1)^\ell A^{00}_{\ell\ell'} \delta_{\ell\ell'} / \sqrt{2\ell+1}$.

A simple model that results in the violation of statistical isotropy arises from the modulation of the of the CMB sky,

$$T(\boldsymbol{n}) = T_0(\boldsymbol{n})\,(1 + M(\boldsymbol{n}))\,, \tag{41}$$

where $T(\boldsymbol{n})$ represents the modulated CMB sky, $T_0(\boldsymbol{n})$ is the underlying statistically isotropic random CMB sky and $M(\boldsymbol{n})$ is a fixed, zero-mean, dimensionless, modulation field. The modulation signal, if any, is expected to be weak and allows quadratic terms in M to be neglected. The BipoSH coefficients for the modulated CMB field ($L > 0$) are then given by the following expression,

$$\begin{aligned} A^{LM}_{\ell_1\ell_2} &= \bar{A}^{LM}_{\ell_1\ell_2} + m_{LM} G^L_{\ell_1\ell_2} \\ G^L_{\ell_1\ell_2} &= \frac{C_{\ell_1} + C_{\ell_2}}{\sqrt{4\pi}} \frac{\Pi_{\ell_1}\Pi_{\ell_2}}{\Pi_L} C^{L0}_{\ell_1 0 \ell_2 0}\,, \end{aligned} \tag{42}$$

where $\bar{A}^{LM}_{\ell_1\ell_2}$ corresponds to the BipoSH coefficients of the unknown, but statistically isotropic, unmodulated CMB field, m_{LM} are the spherical harmonic coefficients of the modulating field ($L > 0$), C_ℓ is the best-fit CMB angular power spectrum and $\Pi_\ell = \sqrt{2\ell+1}$. The statistically isotropic nature of the unmodulated CMB sky implies that the expectation values of $\bar{A}^{LM}_{\ell_1\ell_2}$ vanish for ($L > 0$), leading to the estimator for the modulation field harmonics,

$$\hat{m}_{LM} = \sum_{\ell_1\ell_2} w^L_{\ell_1\ell_2} \frac{\hat{A}^{LM}_{\ell_1\ell_2}}{G^L_{\ell_1\ell_2}}. \tag{43}$$

denoted by the overhat (Hanson & Lewis 2009). The weights $w^L_{\ell_1\ell_2}$ for a minimum variance estimate for the modulation field correspond to

$$w^L_{\ell_1\ell_2} = N \left(\frac{G^L_{\ell_1\ell_2}}{\sigma_{A^{LM}_{\ell_1\ell_2}}} \right)^2 , \tag{44}$$

where N is a normalisation chosen such that $\sum_{\ell_1\ell_2} w^L_{\ell_1\ell_2} = 1$. The BipoSH representation further allows an estimate of the modulation field over specific angular scales by windowing regions in multipole space in the sum over multipoles ℓ_1, ℓ_2 in eqn. 43.

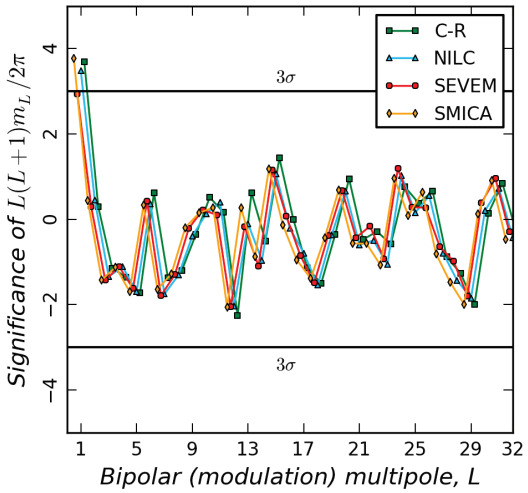


Fig. 32. The significance of the modulation power, $L(L+1)m_L/2\pi$, at bipolar multipoles L. The modulation spectra obtained from the four component separation maps (`C-R`, `NILC`, `SEVEM` and `SMICA`) are consistent with each other. Dipole ($L = 1$) modulation power is detected in all the spectra at a significance ranging from 3.7 to 2.9σ. The solid black lines denote the 3σ significance thresholds. There is no significant power detected at higher multipole of the modulation field $1 < L \leq 32$.

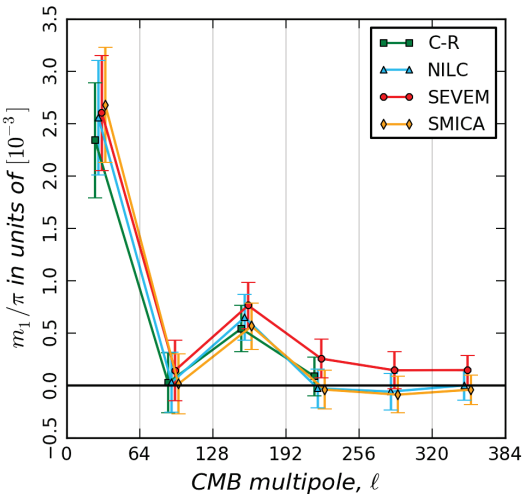


Fig. 33. The CMB multipole dependence of the BipoSH (modulation) power $L(L+1)m_L/2\pi$ can be dissected into bins in ℓ-space. This figure plots the measured dipole modulation ($L = 1$) power in CMB multipole bins. We establish that significant power in the dipole modulation is limited to $\ell \in (2, 64)$ and does not extend to the higher CMB multipoles, ℓ, considered. The vertical grid lines denote the CMB multipole ℓ-bins.

Table 24. This table lists the amplitude and direction of the dipole modulation in Galactic coordinates. The measured values of the dipole amplitude and direction are consistent for all maps. The corresponding dipole power for the `SMICA` map is seen at a detection significance of 3.7σ as shown in Fig. 32.

Map	Dipole Amplitude A	(l,b) [°] ($\sigma_l = 15.4, \sigma_b = 15.1$)
`C-R`	$0.072^{+0.01}_{-0.01}$	(218.9, −21.4)
`NILC`	$0.070^{+0.01}_{-0.01}$	(220.3 , −20.2)
`SEVEM`	$0.065^{+0.011}_{-0.011}$	(221.7 , −21.4)
`SMICA`	$0.073^{+0.01}_{-0.01}$	(217.5 , −20.2)

This additional information could be very useful in identifying the origin of the statistical isotropy violation, which could be either cosmological or due to systematic artefacts (see Hajian & Souradeep 2003; Hajian & Souradeep 2006).

First, we limit our analysis to the four low resolution $N_{\rm side} = 32$ CMB maps used in Sect.5.5.2 and reconstruct the modulation maps for each of them at the same low resolution. The U73 mask is applied to the reconstructed modulation maps before computing m_{LM}. The pseudo-power m_L is corrected for the mask applied to the modulation maps. Specifically for the case of dipole modulation, the pseudo-power m_L is related to the dipole amplitude by $A = 1.5\sqrt{m_1/\pi}$.

A dipole modulation ($L = 1$) signal is detected at 3σ significance in all the maps, as shown in Fig. 32. The amplitude and direction of the dipole modulation match those obtained via a likelihood analysis in Sect. 5.5.2. The BipoSH representation of modulation confirms the dipole modulation signal found in the low-resolution map. Since this approach allows the reconstruction of any general small amplitude modulation field, the BipoSH representation places constraints on the power in the modulation field at all higher (bipolar) multipoles allowed by the resolution of the CMB maps.

We then extend the analysis to higher resolution using maps at $N_{\rm side} = 256$ for `Commander` and $N_{\rm side} = 2048$ for `NILC`, `SEVEM` and `SMICA` in order to study the above effect in more detail. We repeat the analysis on these higher resolution maps using the U73 mask. Contrary to our expectations based on a scale-independent (i.e., no ℓ-dependence) model, the significance of the dipole does not increase in the high resolution maps. We then subdivide the ℓ-range up to $\ell_{\rm max} = 384$ into uniform bins of size $\Delta\ell = 64$. As seen in Fig. 33, we recover the dipole modulation at over $\sim 3\sigma$ significance only for the lowest bin $\ell \in (2, 64)$. This is consistent with the results in Sect. 5.5.2 and the BipoSH analysis on the corresponding low resolution maps shown in Fig. 32. However, the amplitude of the dipole is consistent with zero within 3σ for all of the higher ℓ-bins considered. This suggests that the simple modulation model in Eqn. 41 is inadequate and should minimally allow for the amplitude, $A(\ell)$, of the dipole to depend on CMB multipole, ℓ. Although this may appear to be a more complex model, it does not necessarily lack motivation. It is readily conceivable that physical mechanisms that cause a dipolar modulation of the random CMB sky would be scale dependent and possibly significant only at low wavenumbers. More importantly, such a dipole modulation has also been noted in low resolution *WMAP* data (Eriksen et al. 2007a; Hoftuft et al. 2009). More recently, Bennett et al. (2011) also comment (without being quantitative) that the effect is present in the *WMAP* maps

but limited to low ℓ and conclude that the ℓ dependence rules out a simple modulation explanation. The fact that two independent experiments find this intriguing statistical isotropy violation points to a non-instrumental origin.

It is, of course, possible to extract the BipoSH coefficients $A^{LM}_{\ell_1\ell_2}$, up to the maximum multipole $\ell_{\rm max}$ allowed by the full resolution *Planck* maps at modest computational expense. This allows us to address a specific indication of statistical isotropy violation previously reported in the literature. Bennett et al. (2011) found nonzero BipoSH power spectra, $A^{20}_{\ell\ell}$ and $A^{20}_{\ell\ell+2}$ at very high statistical significance in the *WMAP* maps as determined in ecliptic coordinates, corresponding to a quadrupolar power asymmetry in the CMB sky. The BipoSH spectra peaked at $\ell \sim 250$, and the differences in the BipoSH signal determined from two different frequency bands indicated a non-cosmological origin. Furthermore, the azimuthal symmetry of this BipoSH signal in ecliptic coordinates suggested that it had its origin in some unaccounted-for systematic effect. The findings of Hanson et al. (2010); Joshi et al. (2012) strongly suggest that the signal arises due to an incomplete treatment of beam asymmetries in the data. Bennett et al. (2012) have subsequently noted that analysis of the *WMAP*9 beam-deconvolved maps no longer detects the signal.

We have computed the $A^{20}_{\ell\ell}$ and $A^{20}_{\ell\ell+2}$ in Ecliptic coordinates for the full resolution *Planck* CMB maps as shown in Fig. 34. The analysis yields no evidence for BipoSH coefficients that deviate significantly from zero. This provides conclusive observational evidence from independent CMB measurements that the *WMAP* result could have only arisen due to instrumental artefacts in that data set.

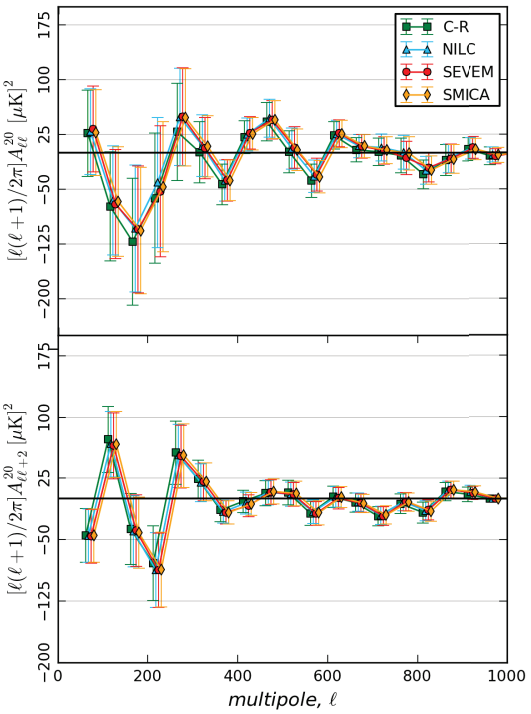


Fig. 34. The BipoSH power spectra $A^{20}_{\ell\ell}$ and $A^{20}_{\ell\ell+2}$ obtained from the four component separation maps (`C-R`, `NILC`, `SEVEM` & `SMICA`) are consistent with each other. Note that no significant ($> 3\sigma$) detections are found. This independently establishes the fact that the quadrupolar BipoSH detections made by *WMAP* were due to *WMAP*-specific instrument systematics.

5.7. Parity asymmetry

5.7.1. Point-parity asymmetry

The CMB sky map may be considered as the sum of even and odd parity functions. Previously, an odd point-parity preference (hereafter parity asymmetry) was observed in the *WMAP* 3-, 5- and 7-year data releases (Land & Magueijo 2005b; Kim & Naselsky 2010a; Naselsky et al. 2012; Kim & Naselsky 2010b; Gruppuso et al. 2011). In this section we investigate the parity asymmetry for the *Planck* temperature anisotropy power spectra derived with a quadratic maximum likelihood (QML) estimator applied to the `Commander-Ruler`, `NILC`, `SEVEM`, and `SMICA` maps at $N_{\rm side} = 32$, and with a pseudo-C_ℓ estimator at $N_{\rm side} = 64$.

From the CMB anisotropy field defined on the sky, $T(\boldsymbol{n})$, we may construct symmetric and antisymmetric functions using the coordinate inversion $\boldsymbol{n} \rightarrow -\boldsymbol{n}$:

$$T^+(\boldsymbol{n}) = \frac{T(\boldsymbol{n}) + T(-\boldsymbol{n})}{2}, \quad T^-(\boldsymbol{n}) = \frac{T(\boldsymbol{n}) - T(-\boldsymbol{n})}{2}. \tag{45}$$

Therefore, $T^+(\boldsymbol{n})$ and $T^-(\boldsymbol{n})$ have even and odd parity, respectively. When combined with the parity property of spherical harmonics, $Y_{\ell m}(\boldsymbol{n}) = (-1)^\ell\, Y_{\ell m}(-\boldsymbol{n})$, we obtain:

$$\begin{aligned} T^+(\boldsymbol{n}) &= \sum_{\ell,m} a_{\ell m}\, Y_{\ell m}(\boldsymbol{n})\, \Gamma^+(\ell), \\ T^-(\boldsymbol{n}) &= \sum_{\ell,m} a_{\ell m}\, Y_{\ell m}(\boldsymbol{n})\, \Gamma^-(\ell), \end{aligned} \tag{46}$$

where n is an integer, and $\Gamma^+(\ell) = \cos^2\left(\frac{\ell\pi}{2}\right)$, and $\Gamma^-(\ell) = \sin^2\left(\frac{\ell\pi}{2}\right)$.

A significant power asymmetry between even and odd multipoles may thus be interpreted as a preference for a particular parity of the anisotropy pattern, connected to the parity asymmetry of the metric perturbations at scales above $1 - 4$ Gpc (Kim & Naselsky 2010a). For investigation of the parity asymmetry we may consider the following quantities (Kim & Naselsky 2010a):

$$\begin{aligned} P^+(\ell) &= \sum_{n=2}^{\ell} \Gamma^+(n) \frac{n(n+1)}{2\pi}\, C_n, \\ P^-(\ell) &= \sum_{n=2}^{\ell} \Gamma^-(n) \frac{n(n+1)}{2\pi}\, C_n, \\ g(\ell) &= \frac{P^+(\ell)}{P^-(\ell)} \end{aligned} \tag{47}$$

where P^+ and P^- are the sum of $n(n+1)/2\pi\, C_n$ for even and odd multipoles respectively; the ratio P^+/P^- is associated with the degree of parity asymmetry, where a value of $P^+/P^- < 1$ indicates odd-parity preference, and $P^+/P^- > 1$ indicates even-parity preference.

Following (Kim & Naselsky 2010a), we will discuss the range of multipoles $2 \le \ell \le 30$, which belongs to the Sachs-Wolfe plateau of the TT power spectrum, where $\ell(\ell+1)C_\ell \sim$ const. In order to make a rigorous assessment of the statistical significance of parity asymmetry at low ℓ, we have compared

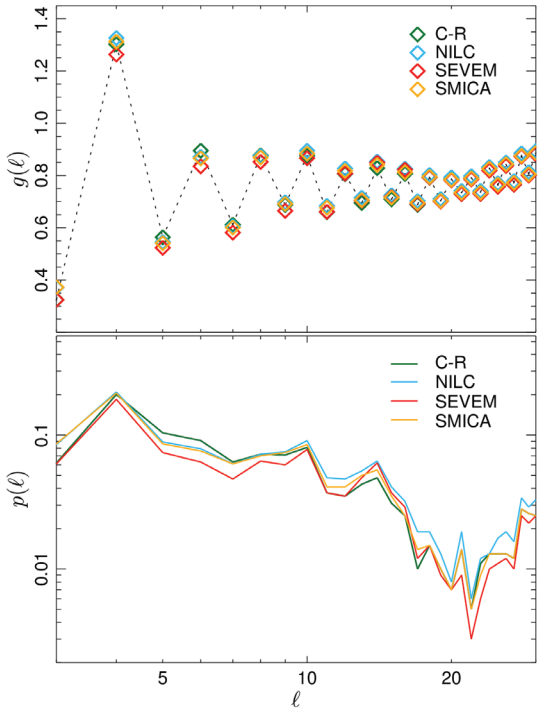


Fig. 35. Top panel: the parity estimator $g(\ell)$ versus ℓ for `Commander-Ruler` (black diamonds), `NILC` (red diamonds), `SEVEM` (blue diamonds) and `SMICA` (green diamonds). Bottom panel: p-value for `C-R` (black solid line), `NILC` (red line), `SEVEM` (blue line), and `SMICA` (green line).

$g(\ell)$ for the *Planck* power spectra with 10^3 simulated CMB maps based on the fiducial *Planck* cosmological model. We compute power spectra using a QML-estimator (Gruppuso et al. 2009) as applied to data at $N_{\rm side} = 32$ with the U73 mask applied. This yields practically identical power spectrum results for the same ℓ-range determined with a pseudo-C_ℓ estimator applied to maps at $N_{\rm side} = 64$.

In Fig.35 we show the $g(\ell)$-parameter for the *Planck* power spectra and the corresponding p-values. The p-value denotes the fraction of simulations in which the obtained value of P^+/P^- is as low as that observed in the data. Note that the results from the different *Planck* CMB maps yield consistent shapes for the $g(\ell)$ and $p(\ell)$-parameters. The parity asymmetry at $\ell = 22$ is most anomalous, with a corresponding p-value in the range $0.002 - -0.004$). Finally, the statistical significance of the parity asymmetry (i.e., low p-value) increases when we increase $\ell_{\rm max}$ up to 22-25. Therefore, the odd parity preference cannot simply be attributed to the low quadrupole power. It is plausible the low quadrupole power is not an isolated anomaly, but that it shares an origin with the odd parity preference (see for details (Kim & Naselsky 2010a; Naselsky et al. 2012; Kim & Naselsky 2010b)).

5.7.2. Mirror Parity

In this section we investigate the properties of the *Planck* temperature low-resolution maps under reflection with respect to a plane. This search for hidden mirror symmetries and anti-symmetries complements the tests for parity asymmetry, presented in Sect. 5.7.1. Starobinsky (1993) showed how a hidden mirror symmetry might be connected to the non-compact T^1 topology, or to a compact T^3 topology in which one topological scale is much less than the others. The CMB pattern would then exhibit a mirror symmetry with respect to the plane defined by the two large dimensions. Mirror symmetry has been searched for in the *COBE*-DMR data in de Oliveira-Costa et al. (1996), resulting in a lower limit for the scale of the compact dimension as 4 Gpc (see also Gurzadyan et al. 2007; Ben-David et al. 2012 for other more recent analysis). Finelli et al. (2012) analysed hidden mirror symmetry and anti-symmetry properties of the *WMAP* 7-year `ILC` temperature map, finding a preferred direction that could be considerede anomalous at the 93 % confidence level with anti-symmetry properties. This direction lies close to the one defining the hemispherical asymmetry.

Following Finelli et al. (2012), we consider the following estimators:

$$S^{\pm}(\boldsymbol{n}_i) = \frac{1}{N_{\rm pix}} \sum_{j=1}^{N_{\rm pix}} \left[\frac{1}{2} \left(\frac{\delta T}{T}(\boldsymbol{n}_j) \pm \frac{\delta T}{T}(\boldsymbol{n}_k) \right) \right]^2 , \quad (48)$$

where the sum is meant over the observed pixels, $N_{\rm pix}$, $\delta T/T(\boldsymbol{n}_j)$ is the CMB temperature anisotropy measured at the pixel pointed by the unit vector $\boldsymbol{n}_j$, and $\boldsymbol{n}_k$ is the opposite direction of $\boldsymbol{n}_j$ with respect to the plane defined by $\boldsymbol{n}_i$, i.e.

$$\boldsymbol{n}_k = \boldsymbol{n}_j - 2\,(\boldsymbol{n}_i \cdot \boldsymbol{n}_j)\boldsymbol{n}_i\,. \quad (49)$$

We compute the quantities $S^{\pm}$ for each of the 3072 directions defined by `HEALPix` resolution $N_{\rm side} = 16$ map, by allowing the j and k indices to run over the unmasked pixels of the low resolution foreground cleaned maps. We perform the same analysis on 1000 simulated skies and store the minimum and maximum value for each of these.

The minimum value for the S^+ estimator is reached for the plane defined by Galactic coordinates $(\theta, \phi) = (104^\circ, 262^\circ)$, with a significance of 0.8% (`Commander-Ruler`), 0.5% (`NILC`), 9.6% (`SEVEM`), and 1.2% (`SMICA`). The top panel of Fig. 36 shows the minimum value of S^+ for each of the four methods and compared to the MC simulations computed for `Commander-Ruler`, which is considered to be representative.

The minimum value for the S^- estimator is found for a direction close to that associated with the cosmological dipole. It is not statistically significant for any of the CMB maps (see bottom panel of Fig.(36)).

The anomalous anti-symmetry direction found in the *Planck* CMB data is close to that found for the dipolar modulation in Sect. 5.5 suggesting some connection between them. The directions which minimize S^+ and S^- for *Planck* are the same as those found for the *WMAP* 7-year `ILC` map in Finelli et al. (2012).

5.8. The Cold Spot

The Cold Spot was identified in the *WMAP* first year data (Vielva et al. 2004) through the estimation of the kurtosis of the Spherical Mexican Hat Wavelet (SMHW, e.g., Martínez-González et al. 2002) coefficients, and confirmed (Cruz et al. 2005) by analysing the area of the SMHW coefficients

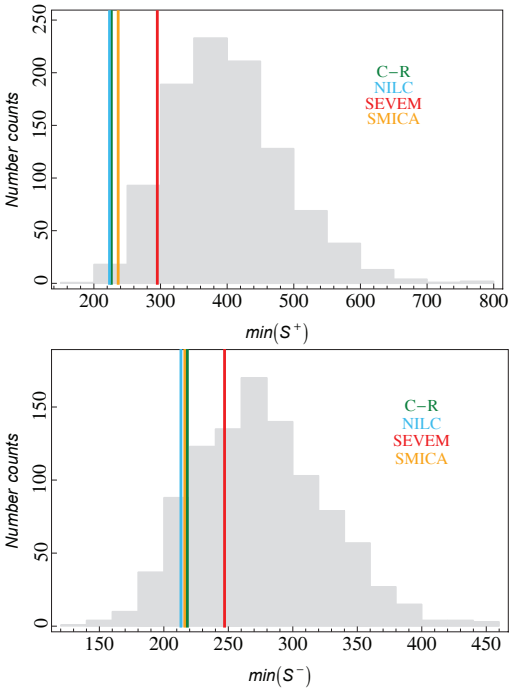


Fig. 36. Top panel: the S^+ statistic. The vertical lines show the minimum value for the estimator as computed on low resolution C-R, NILC, SEVEM, and SMICA maps. The grey histogram shows the same quantity computed from 1000 simulated maps processed by C-R. Bottom panel: as above for S^-.

above/below a given threshold. Since its detection, the Cold Spot has been extensively studied and verified with a large battery of statistical probes (e.g., Mukherjee & Wang 2004; Cayón et al. 2005; McEwen et al. 2005; Cruz et al. 2007a; Räth et al. 2007b; Vielva et al. 2007; Pietrobon et al. 2008; Gurzadyan et al. 2009; Rossmanith et al. 2009b). A complete review of the Cold Spot can be found in Vielva (2010), including a discussion on possible explanations of its nature.

The analysis of the kurtosis of the SMHW coefficients has already been addressed in Sect. 4.5. We have checked that the kurtosis of the coefficients corresponding to the four *Planck* cleaned frequency maps is larger than the expected value obtained from simulations, with a modified upper tail probability of around 0.01. This is compatible with the value obtained from *WMAP*.

Nevertheless, the Cold Spot is more robustly described in terms of a morphological quantity: the area of the SMHW coefficients above/below a given threshold. At a given scale R and threshold ν, the cold ($A_R^{-\nu}$) and hot ($A_R^{+\nu}$) areas of the SMHW coefficients are defined as:

$$A_R^{-\nu} \equiv \#\{\omega_T(R,p) < -\nu\} \qquad (50)$$
$$A_R^{+\nu} \equiv \#\{\omega_T(R,p) < +\nu\}$$

where # represents the number operator, i.e, it indicates for how many pixels p, the specific condition defined between the braces is satisfied.

Table 25. Upper tail probability (UTP, in %) associated to the cold (left) and hot (right) areas. Results are given for a $\nu > 4\sigma_R$ threshold and for the four *Planck* CMB maps. The three most significant scales associated to the Cold Spot are shown. Analysis performed on the exclusions masks associated with the U73 mask.

	Scale (R) [′]	C-R UTP (%)	NILC UTP (%)	SEVEM UTP (%)	SMICA UTP (%)
cold area	200	1.6	1.1	1.2	1.1
	250	0.3	0.3	0.3	0.3
	300	0.3	0.3	0.3	0.3
hot area	200	2.3	1.6	1.8	1.6
	250	2.7	2.2	2.4	2.2
	300	4.9	3.7	4.1	3.8

Table 26. Upper tail probability (in %) associated to the cold (left) and hot (right) areas. Results are given for a $\nu > 4\sigma_R$ threshold and for the four *Planck* CMB maps. The three most significant scales associated to the Cold Spot are shown. Analysis performed on the exclusions masks associated with the G70 mask. N/A indicates that no area above that threshold was found on the data.

	Scale (R) [′]	C-R UTP (%)	NILC UTP (%)	SEVEM UTP (%)	SMICA UTP (%)
cold area	200	1.0	1.0	0.9	1.0
	250	0.3	0.3	0.3	0.3
	300	0.2	0.2	0.3	0.2
hot area	200	15.1	14.5	14.6	14.5
	250	N/A	N/A	N/A	N/A
	300	N/A	N/A	N/A	N/A

Table 25 summarises the results for the hot and cold areas determined for the four CMB maps analysed with the U73 mask (and its associated exclusions masks). The cold area is anomalous at scales between $R = 200$ and $R = 300'$, similar to the sizes already highlighted with the kurtosis analysis. We see that the higher the threshold, the smaller the upper tail probability associated with the *Planck* CMB map. In particular, the cold area has a upper tail probability of 0.003 at $\nu > 4\sigma_R$ and for $R = 300'$.

Notice that the most significant deviation comes from the cold area, although the hot area is marginally compatible. However, the cold area represents the most robust detection of an anomaly, since it is robust to the mask employed (see Tables 26 and 27).

The information provided in the previous Tables is also represented (for the $R = 300'$scale) in Fig. 37. In these nine panels we show the anomalous cold (in blue) and hot (in red) areas for thresholds $\nu > 3.0\sigma_R$, $\nu > 3.5\sigma_R$ and $\nu > 4.0\sigma_R$ as determined from the SMICA map. For the two largest thresholds, the cold area corresponds to the Cold Spot, whereas the red area at 3.0σ has already been identified in the *WMAP* data (e.g., Vielva et al. 2007) as an anomalous hot spot. From these analyses it is clear that the Cold Spot anomaly is present in both the *WMAP* and *Planck* data.

5.9. Interpretation of anomalies

The results presented here in Sect. 5 demonstrate that many features previously observed in the *WMAP* data are present also in

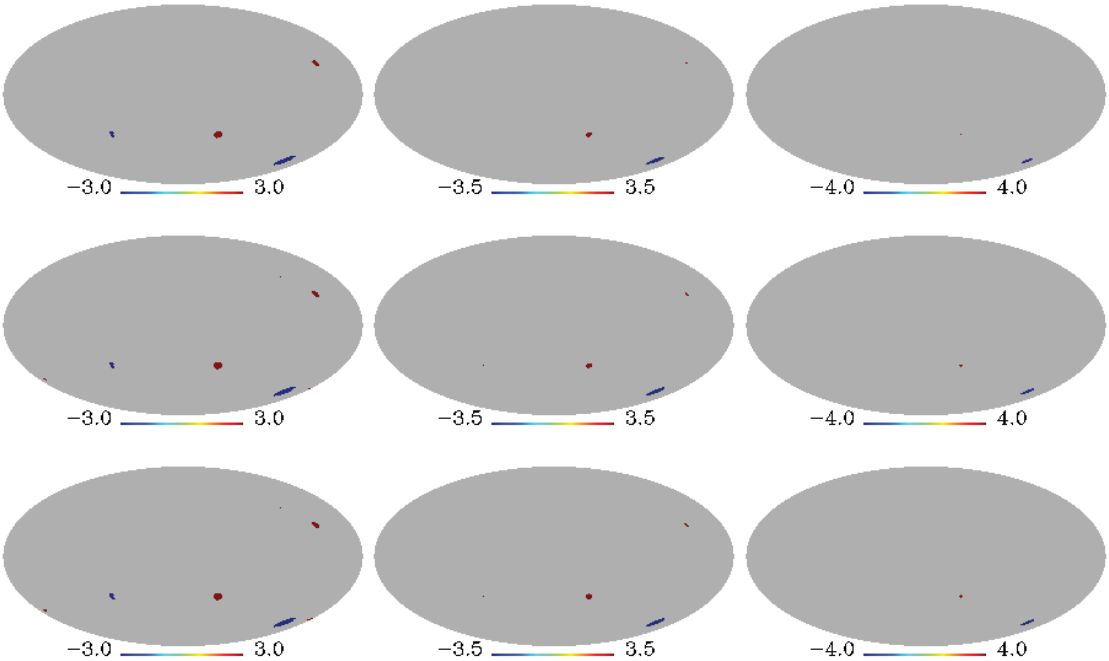


Fig. 37. SMHW coefficients at $R = 300$ arc minutes, and thresholds of 3.0σ (left), 3.5σ (middle), and 4.0σ (right). Results for the three masks considered in the analysis are shown: U73 mask (top), CG70 (middle) and CG60 (bottom).

Table 27. Upper tail probabilities (in %) associated with the cold (left) and hot (right) areas. Results are given for a $\nu > 4\sigma_R$ threshold and for the four *Planck* CMB maps. The three most significant scales associated to the Cold Spot are shown. Analysis performed on the exclusions masks associated with the CG60 mask. N/A indicates that no area above that threshold was found on the data.

	Scale (R) [′]	C-R UTP (%)	NILC UTP (%)	SEVEM UTP (%)	SMICA UTP (%)
cold area	200	1.1	0.9	0.8	0.9
	250	0.1	0.1	0.1	0.1
	300	0.1	0.1	0.1	0.1
hot area	200	N/A	N/A	N/A	N/A
	250	N/A	N/A	N/A	N/A
	300	N/A	N/A	N/A	N/A

the *Planck* sky. This agreement between two independent experiments effectively rules out the possibility that their origin lies in systematic artefacts present in either data set. In particular, there is evidence for a violation of statistical isotropy at least on large angular scales in the context of the *Planck* fiducial sky model. Moreover, a dipolar power asymmetry may extend to scales corresponding to $\ell \simeq 1500$, whilst fits to a model containing a dipole modulation field yield results in excess of 3σ significance. In addition, there is evidence from such fits that the low-ℓ spectrum of the *Planck* data departs from the fiducial spectrum in both amplitude and slope. These results could have profound implications for cosmology. It is therefore pertinent to consider whether a model can be proposed to provide a common origin for the anomalies.

The microwave sky is manifestly non-Gaussian and anisotropic, with known contributions from Galactic astrophysical foregrounds, lensing of CMB anistropies by the intervening matter distribution, and the ISW. However, the excellent performances of the component separation algorithms used here in rejecting diffuse foregrounds argues strongly against known Galactic emission as the source of the anomalies.

Schwarz et al. (2004), Copi et al. (2007), Maris et al. (2011) and Hansen et al. (2012) suggested that diffuse Solar System emission could contribute to the observed structure on large angular scales, although it is not expected that the classical Zodiacal Light Emission or Kuiper Belt objects are responsible. Planck Collaboration XIV (2013) presents the current *Planck* contribution to the modelling of the Zodiacal cloud.

Another possibility is that the anomalies have their origin in the local Universe. According to Francis & Peacock (2009), the removal of the ISW signal originating within the volume at $z < 0.3$ from *WMAP* data reduces the significance of the apparent alignment between the CMB quadrupole and octopole and the Cold Spot. Efstathiou et al. (2010) have used the same correction to yield an increase in the structure of the two-point correlation function for angular separations less than 60°, that had been noted as apparently anomalous since the first *WMAP* data release. A future possibility is that *Planck* itself will be able to reconstruct the ISW signal and test its impact on issues related to

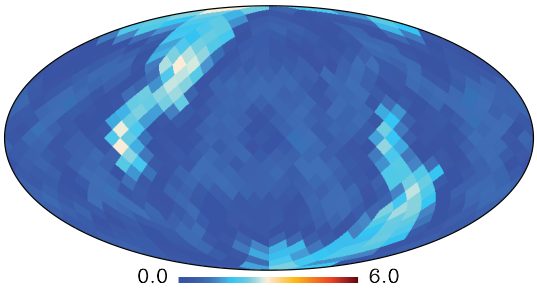


Fig. 38. Same as Fig. 24 but with the best fit Bianchi template subtracted from the SMICA map.

isotropy and non-Gaussianity. Planck Collaboration XIX (2013) presents maps of the effect based on the current data release.

Of more interest to us is that the anomalies are genuinely cosmological in origin. In that context, obvious candidate models include those with simply or multi-connected topology. In a companion paper (Planck Collaboration XXVI 2013), a subset of such models are considered and the signatures of their specific correlation structures on the sky are searched for. However, no detections are found, but rather the scale of topology is limited to be of order the diameter of the last-scattering surface or greater. More interestingly, they reconsider Bianchi $\mathrm{VII_h}$ models that were previously demonstrated to show statistical correlation with the *WMAP* data (Jaffe et al. 2005, 2006; Bridges et al. 2007; McEwen et al. 2013), albeit with parameters inconsistent with standard cosmological parameters. In this new analysis, the Bianchi parameters are physically coupled to the cosmological ones, yielding no evidence for a Bianchi VII_h cosmology. However, as before, when treated simply as a template for structure contained in the CMB sky, a best-fit pattern is found to be in good agreement with the old results. Previous analyses (Jaffe et al. 2005; Cayón et al. 2006; McEwen et al. 2006) have shown that when the CMB sky is corrected for such a template, many of the large-scale anomalies are no longer present at a statistically significant level. It is likely that such an effect will persist for Bianchi-corrected *Planck* data, and we have made an explicit test as to whether the best-fit Bianchi template can also explain the presence of phase correlations. We therefore repeated the surrogate analysis from Sect. 5.4 for the appropriately corrected SMICA map. Figure 38 presents the result for the corresponding significance map. It is clear that the signature for hemispherical asymmetry is drastically reduced, thereby rendering the signal formally statistically insignificant. Thus, the best-fit Bianchi model can also account for the asymmetries induced by higher order phase correlations. It should also be noted that subtracting the best-fit Bianchi template from the data, outside the U73 mask, explains the anomalous skewness and kurtosis values but not the variance, for which the corresponding lower tail probabilities are 0.008, 0.166, and 0.306, respectively. Given the lack of consistency of the physical parameters of the model with the *Planck* cosmological model, the results obtained using Bianchi-subtracted input maps might be considered moot, however, the morphology of the maps may provide insight into the type of underlying structures associated with the anomalies.

Although the Cold Spot is also rendered statistically insignificant by the Bianchi template, other possible explanations about its nature have been proposed, including the late evolution of the large-scale structure (e.g., Inoue & Silk 2006, 2007), the Sunyaev-Zeldovich effect (e.g., Cruz et al. 2008), residual foregrounds (Cruz et al. 2006), gravitational lensing (Das & Spergel 2009), or a cosmic texture (e.g., Cruz et al. 2007b).

The presence of primordial magnetic fields (PMFs) due to either pre- or post-recombination mechanisms could also provide a physical basis for some of the anomalies discussed in this paper. Specifically, PMFs with coherence scales comparable to the present day horizon could result in Alfvén waves in the early Universe that generate specific signatures on the sky via the Doppler and integrated Sachs-Wolfe effects. In particular, a preferred angular direction in the CMB anisotropy can be induced (Durrer et al. 1998; Kim & Naselsky 2009) that leads to structure in the spherical harmonic mode correlation matrix (Kahniashvili et al. 2008). Appendix A presents a search for the predicted correlations between harmonic modes separated by $\Delta\ell = 0, \pm2$, and $\Delta m = 0, \pm2$, allowing constraints to be placed on the Alfvén wave amplitude. Further constraints on PMFs based on the power spectrum and bispectrum have been provided in Planck Collaboration XVI (2013) and Planck Collaboration XXIV (2013), respectively.

To conclude, when analysing a data set as complex and rich as that provided by *Planck*, some statistical outliers will be expected. However, it should be clear that the evidence for some of the large-angular scale anomalies is significant indeed, yet few physically compelling models have been proposed to account for them, and none so far that provide a common origin. The dipole modulation model that was analysed here was phenomenologically motivated, and is detected in the data at relatively high significance. Whether it can resolve the anomalous nature of other observed features remains to be evaluated.

6. Implications for C_ℓ and cosmological parameter estimation

The approach to C_ℓ estimation, the construction of the *Planck* likelihood and subsequent inference of cosmological parameters are described in the accompanying papers Planck Collaboration XV (2013); Planck Collaboration XVI (2013). For these studies, specific assumptions are made about the isotropy and Gaussianity of the primordial fluctuations observed in the CMB. The latter in particular seems to be well-supported by the comprehensive set of tests applied to the *Planck* data in Sect. 4. The most significant discrepancies are seen in association with the Cold Spot (Sect. 5.8), which constitutes a localized region of sky of about 10° in size. Its impact on cosmological parameters is then likely to be relatively insignificant, and masking of the region could easily test this assertion.

It is well-known that the quadrupole and octopole have low-amplitudes relative to the best-fit cosmological power-spectrum. The contribution of those multipoles to cosmological parameter estimation is very small due to the associated cosmic variance on these scales, and thus the direct impact of their alignment (as discussed in Sect. 5.1) is also likely to be small. Remarkably, however, Planck Collaboration XV (2013) presents evidence that the low-ℓ multipole range from 2-30 is coherently low, and is not well accounted for by the standard ΛCDM model. Moreover, this conclusion is a consequence of the fact that the cosmological parameters are strongly influenced by the ℓ =1000–1500 range, previously inaccessible to *WMAP*. Consistent findings have been presented here in the form of the low-variance of the data in Sects. 4.1 and 5.2, although this is largely driven by the quadupole and octopole alone.

The question therefore remains as to whether there is a deeper connection with the cosmological anomalies seen in both the *WMAP* and *Planck* data sets particularly on large angular scales. Indeed, the hemispherical asymmetry and dipolar power modulation discussed in Sects. 5.3 and 5.5 respectively could have a more important impact in that they directly address whether a broader class of cosmological models should be considered. Indeed, the low-ℓ signature seen in the data has previously been associated with missing power in a Universe with simply- or multiply-connected topology. However, there are specific morphological signatures of such topologies that have not been detected in the *Planck* data (Planck Collaboration XXVI 2013).

However, the phenomenologically motivated dipole modulation model due to Gordon et al. (2005) yields a significant fit to the data, as seen in Sects. 5.5.2 and 5.6. The former also shows some evidence for a departure from the *Planck* fiducial power spectrum in both amplitude and slope. Both of these analyses are in good agreement in terms of the direction of the dipolar modulation field with the model independent dipolar power modulation analysis of Sect. 5.5.1.

A qualitative exploration as to how these features are reflected in the low-ℓ power spectrum is provided in Fig. 39. Specifically, the plot presents the angular power spectra computed using a quadratic maximum-likelihood (QML) estimator (Paci et al. 2010, 2013) from the $N_{\rm side} = 16$ SMICA map after application of the U73 mask used in this paper. The *Planck* fiducial power spectrum is also shown for comparison. Clearly, there is a deficit of power as expected when no further partitioning of the sky is applied. However, further interesting properties of the data are revealed when spectra are computed for the two opposing hemispheres defined by the preferred direction in Sect. 5.5.1. In the positive direction, there is improved agreement between the derived spectrum and the *Planck* fiducial sky, but with an interesting oscillation between odd and even modes. For the negative direction, an overall suppression of power is again seen. It would be interesting to test the connection between these spectral features and the phase correlations detected in Sect. 5.4 or the evidence for parity violation presented in Sect. 5.7. The observations may, in part, reflect the presence of visually striking features noted by Bennett et al. (2011) — the four elongated cold fingers stretching from near the Galactic equator to the south Galactic pole and a prominent cold spot near the center of the map.

However, Fig. 28 and the corresponding analysis suggest that the asymmetry in power between hemispheres extends to much smaller angular scales. Whether such a property of the data would have implications for parameter estimation may yet need further exploration.

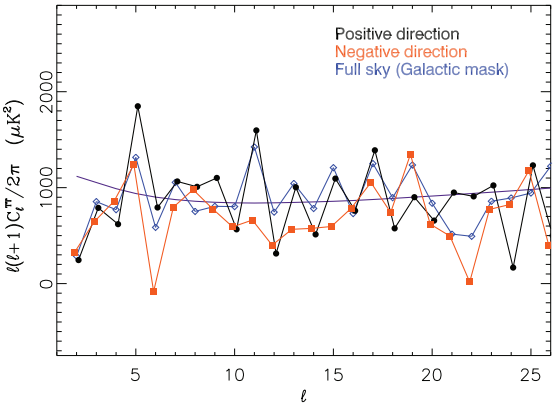


Fig. 39. Angular power spectrum at large angular scales computed on opposing hemispheres defined by the maximal asymmetry.

7. Conclusions

In this paper, we have tested the statistical isotropy and Gaussianity of the CMB using data from the *Planck* satellite. We have demonstrated that little evidence is seen for non-Gaussianity, although some deviations from isotropy are found.

Most of the tests performed in Sect. 4 showed an overall consistency with the null hypothesis, as represented by a set of realistic simulations based on a *Planck* fiducial sky model and including the secondary ISW-lensing effect (as detected for the first time with the *Planck* data, see Planck Collaboration XIX 2013). However, two important exceptions were seen. The variance of the CMB signal was found to be significantly lower than expected, with the anomalously low signal seemingly localized in the northern ecliptic hemisphere (Sect. 4.1). This result was also confirmed with the low variance of the wavelet coefficients that was seen on scales above a few degrees (see Sect. 4.5). Moreover, a significant deviation from Gaussianity was found in the form of a positive kurtosis of the wavelet coefficients.

These results correspond to statistical features on large angular scales where numerous anomalies were previously observed in the *WMAP* data. In Sect. 5, we revisited these in the light of the *Planck* data and found results in excellent agreement with those for *WMAP*. In particular, the most significant anomalies, namely the quadrupole-octopole alignment (Sect. 5.1), hemispherical asymmetry (Sect. 5.3), phase correlations (Sect. 5.4), dipolar power modulation (Sect. 5.5), generalized power modulation (Sect. 5.6), parity asymmetry (Sect. 5.7) and the Cold Spot (Sect. 5.8) have been confirmed with the *Planck* data. Attempts to explain the observed features in terms of systematic artefacts, local astrophysical sources of emission, or structure in the local Universe have not been successful. It is clear that these anomalies represent real features of the CMB sky.

However, it is difficult to make a detailed interpretation of the anomalies in the absence of theoretical models, in particular with regard to the role of a posterior choices. Nevertheless, *Planck* does offer new possibilities to check the a posteriori claims in this context as a consequence of its superior multipole content that cannot easily been probed by *WMAP*. This is particularly relevant for the power asymmetry studies — Sect. 5.5.1 found that the same direction was preferred at $\ell > 600$ as for $\ell < 50$, which should mitigate in part such criticisms.

Phenomenological models have been suggested to account for the observations. The dipolar power modulation approach due to Gordon et al. (2005) was explicitly tested in Sect. 5.5.2 and found to represent a good fit to the large scale asymmetry, corresponding to a detection at about 3σ significance. This result was confirmed by the more generalized modulation study in Sect. 5.6, which also ruled out the presence of modulation fields of higher order. Alternatively, a Bianchi template fit to the data performed in Planck Collaboration XXVI (2013) can provide a good fit to the hemispherical asymmetry, the Cold Spot and the phase correlations, but corresponds to values of the cosmological parameters incompatible with those derived in

Planck Collaboration XVI (2013). Clearly, these do not provide complete and satisfactory explanations for the observations, and more physically motivated models should be sought.

This may also be indicated by the cosmological parameter studies presented in Planck Collaboration XV (2013); Planck Collaboration XVI (2013). Here, it was demonstrated that while the power-spectrum determined from the *Planck* temperature data is extremely consistent with a basic 6-parameter ΛCDM model, the low-ℓ multipoles ($\ell \leq 30$) deviate from the best-fit model although at a significance that does not appear to exceed 2.7σ. However, this is precisely the regime where many of the anomalies presented in this paper seem to manifest themselves, and where qualitatively interesting differences are observed in the power-spectra for two hemispheres defined by the preferred direction for the dipolar power modulation.

Finally, it is expected that the polarization data that will become available with the 2014 data release should provide valuable information on the nature of the CMB anomalies. Then, the presence, or even absence, of a specific signature in the data should help to elucidate the physical mechanism that is causing the anomaly (see Vielva et al. 2011, Frommert & Enßlin 2010 and Dvorkin et al. 2008 for examples related to the Cold Spot, mode alignment, and dipolar modulation, respectively) In particular, a deviation of isotropy present at recombination should be reflected in both the temperature and polarization data with a correlated signal. It may be that the statistical anomalies currently described in this paper are a hint of more profound physical phenomena that are yet to be revealed.

Acknowledgements. The development of Planck has been supported by: ESA; CNES and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MICINN, JA and RES (Spain); Tekes, AoF and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); and PRACE (EU). A description of the Planck Collaboration and a list of its members, including the technical or scientific activities in which they have been involved, can be found at `http://www.sciops.esa.int/index.php?project=planck&page=Planck_Collaboration`. We acknowledge the use of resources from the Norewegian national super computing facilities NOTUR. The modal and KSW bispectrum estimator analysis was performed on the COSMOS supercomputer, part of the STFC DiRAC HPC Facility. We further acknowledge the computer resources and technical assistance provided by the Spanish Supercomputing Network nodes at Universidad de Cantabria and Universidad Politécnica de Madrid as well as by the Advanced Computing and e-Science team at IFCA.

References

Ben-David, A., Kovetz, E. D., & Itzhaki, N. 2012, Astrophys.J., 748, 39
Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2011, ApJS, 192, 17
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2012, ArXiv e-prints
Bielewicz, P., Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2005, ApJ, 635, 750
Bielewicz, P., Górski, K. M., & Banday, A. J. 2004, MNRAS, 355, 1283
Bielewicz, P. & Riazuelo, A. 2009, MNRAS, 396, 609
Bielewicz, P., Wandelt, B. D., & Banday, A. J. 2013, MNRAS, 429, 1376
Bridges, M., McEwen, J. D., Lasenby, A. N., & Hobson, M. P. 2007, MNRAS, 377, 1473
Cayón, L., Banday, A. J., Jaffe, T., et al. 2006, MNRAS, 369, 598
Cayón, L., Jin, J., & Treaster, A. 2005, MNRAS, 362, 826
Copi, C. J., Huterer, D., Schwarz, D. J., & Starkman, G. D. 2007, Phys.Rev.D, 75, 023507
Copi, C. J., Huterer, D., & Starkman, G. D. 2004, Phys. Rev. D, 70, 043515
Cruz, M., Cayón, L., Martínez-González, E., Vielva, P., & Jin, J. 2007a, ApJ, 655, 11
Cruz, M., Martínez-González, E., Vielva, P., & Cayón, L. 2005, MNRAS, 356, 29
Cruz, M., Martínez-González, E., Vielva, P., et al. 2008, MNRAS, 390, 913
Cruz, M., Tucci, M., Martínez-González, E., & Vielva, P. 2006, MNRAS, 369, 57
Cruz, M., Turok, N., Vielva, P., Martínez-González, E., & Hobson, M. 2007b, Science, 318, 1612
Cruz, M., Vielva, P., Martínez-González, E., & Barreiro, R. B. 2011, MNRAS, 412, 2383
Curto, A., Aumont, J., Macías-Pérez, J. F., et al. 2007, A&A, 474, 23
Curto, A., Macías-Pérez, J. F., Martínez-González, E., et al. 2008, A&A, 486, 383
Curto, A., Martínez-González, E., & Barreiro, R. B. 2009a, ApJ, 706, 399
Curto, A., Martínez-González, E., & Barreiro, R. B. 2010, in Highlights of Spanish Astrophysics V, ed. J. M. Diego, L. J. Goicoechea, J. I. González-Serrano, & J. Gorgas, 277–+
Curto, A., Martínez-González, E., & Barreiro, R. B. 2011a, MNRAS, 412, 1038
Curto, A., Martínez-González, E., Barreiro, R. B., & Hobson, M. P. 2011b, MNRAS, 417, 488
Curto, A., Martínez-González, E., Mukherjee, P., et al. 2009b, MNRAS, 393, 615
Das, S. & Spergel, D. N. 2009, Phys. Rev. D, 79, 043007
de Oliveira-Costa, A., Smoot, G. F., & Starobinsky, A. A. 1996, Astrophys.J., 468, 457
de Oliveira-Costa, A., Tegmark, M., Zaldarriaga, M., & Hamilton, A. 2004, Phys.Rev.D, 69, 063516
De Troia, G., Ade, P. A. R., Bock, J. J., et al. 2007, ApJ, 670, L73
Ducout, A., Bouchet, F. R., Colombi, S., Pogosyan, D., & Prunet, S. 2012, MNRAS, 423
Durrer, R., Kahniashvili, T., & Yates, A. 1998, Phys. Rev. D, 58, 123004
Dvorkin, C., Peiris, H. V., & Hu, W. 2008, Phys. Rev. D, 77, 063008
Efstathiou, G. 2004, MNRAS, 348, 885
Efstathiou, G., Ma, Y.-Z., & Hanson, D. 2010, MNRAS, 407, 2530
Elsner, F. & Wandelt, B. D. 2013, A&A, 549, A111
Eriksen, H. K., Banday, A. J., Górski, K. M., Hansen, F. K., & Lilje, P. B. 2007a, ApJ, 660, L81
Eriksen, H. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2005, ApJ, 622, 58
Eriksen, H. K., Hansen, F. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004a, ApJ, 605, 14
Eriksen, H. K., Huey, G., Saha, R., et al. 2007b, ApJ, 656, 641
Eriksen, H. K., Lilje, P. B., Banday, A. J., & Górski, K. M. 2004b, ApJS, 151, 1
Eriksen, H. K., Novikov, D. I., Lilje, P. B., Banday, A. J., & Górski, K. M. 2004c, ApJ, 612, 64
Fergusson, J., Liguori, M., & Shellard, E. 2010a, Phys.Rev., D82, 023502
Fergusson, J., Liguori, M., & Shellard, E. 2010b
Finelli, F., Gruppuso, A., Paci, F., & Starobinsky, A. 2012, JCAP, 1207, 049
Fixsen, D. J. 2009, ApJ, 707, 916
Francis, C. L. & Peacock, J. A. 2009, MNRAS, 406,, 14
Frommert, M. & Enßlin, T. A. 2010, MNRAS, 403, 1739
Gordon, C., Hu, W., Huterer, D., & Crawford, T. 2005, Phys.Rev.D, 72, 103002
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Gott, J. R. I., Park, C., Juszkiewicz, R., et al. 1990, ApJ, 352, 1
Gruppuso, A., de Rosa, A., Cabella, P., et al. 2009, MNRAS, 400, 463
Gruppuso, A., Finelli, F., Natoli, P., et al. 2011, MNRAS, 411, 1445
Gurzadyan, V., Starobinsky, A., Kashin, A., Khachatryan, H., & Yegorian, G. 2007, Mod.Phys.Lett., A22, 2955
Gurzadyan, V. G., Allahverdyan, A. E., Ghahramanyan, T., et al. 2009, A&A, 497, 343
Hajian, A. & Souradeep, T. 2003, ApJ, 597, L5

Hajian, A. & Souradeep, T. 2006, Phys. Rev. D, 74, 123521
Hansen, F. K., Banday, A. J., & Górski, K. M. 2004, MNRAS, 354, 641
Hansen, F. K., Banday, A. J., Górski, K. M., Eriksen, H. K., & Lilje, P. B. 2009, ApJ, 704, 1448
Hansen, M., Kim, J., Frejsel, A. M., et al. 2012, J. Cosmology Astropart. Phys., 10, 59
Hanson, D. & Lewis, A. 2009, Phys.Rev.D, 80:063004,2009
Hanson, D., Lewis, A., & Challinor, A. 2010, Phys. Rev. D, 81, 103003
Hikage, C., Matsubara, T., Coles, P., et al. 2008, MNRAS, 389, 1439
Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2
Hoftuft, J., Eriksen, H. K., Banday, A. J., et al. 2009, Astrophys.J., 699:985-989,2009
Inoue, K. T. & Silk, J. 2006, ApJ, 648, 23
Inoue, K. T. & Silk, J. 2007, ApJ, 664, 650
Jaffe, T. R., Banday, A. J., Eriksen, H. K., Górski, K. M., & Hansen, F. K. 2005, ApJ, 629, L1
Jaffe, T. R., Hervik, S., Banday, A. J., & Górski, K. M. 2006, ApJ, 644, 701
Joshi, N., Das, S., Rotti, A., Mitra, S., & Souradeep, T. 2012, ArXiv e-prints
Kahniashvili, T., Lavrelashvili, G., & Ratra, B. 2008, Phys. Rev. D, 78, 063012
Kim, J. & Naselsky, P. 2009, J. Cosmology Astropart. Phys., 7, 41
Kim, J. & Naselsky, P. 2010a, ApJ, 714, L265
Kim, J. & Naselsky, P. 2010b, Phys. Rev. D, 82, 063002
Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2009, ApJS, 180, 330
Komatsu, E., Kogut, A., Nolta, M. R., et al. 2003, ApJS, 148, 119
Kronberg, P. P. 2009, in IAU Symposium, Vol. 259, IAU Symposium, 499–508
Land, K. & Magueijo, J. 2005a, Physical Review Letters, 95, 071301
Land, K. & Magueijo, J. 2005b, Phys. Rev. D, 72, 101302
Lewis, A. & Bridle, S. 2002, Phys. Rev. D, 66, 103511
Maris, M., Burigana, C., Gruppuso, A., Finelli, F., & Diego, J. M. 2011, MNRAS, 415, 2546
Martínez-González, E., Gallegos, J. E., Argüeso, F., Cayón, L., & Sanz, J. L. 2002, MNRAS, 336, 22
Matsubara, T. 2010, Phys. Rev. D, 81, 083505
McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2005, MNRAS, 359, 1583
McEwen, J. D., Hobson, M. P., Lasenby, A. N., & Mortlock, D. J. 2006, MNRAS, 369, 1858
McEwen, J. D., Josset, T., Feeney, S. M., Peiris, H. V., & Lasenby, A. N. 2013, ArXiv e-prints
Mecke, K. R., Buchert, T., & Wagner, H. 1994, A&A, 288, 697
Minkowski, H. 1903, Mathematische Ann., 57, 447
Mitra, S., Rocha, G., Górski, K. M., et al. 2011, ApJS, 193, 5
Modest, H. I., Räth, C., Banday, A. J., et al. 2013, MNRAS, 428, 551
Monteserín, C., Barreiro, R. B., Vielva, P., et al. 2008, MNRAS, 387, 209
Moss, A., Scott, D., Zibin, J. P., & Battye, R. 2011, Phys. Rev. D, 84, 023014
Mukherjee, P. & Wang, Y. 2004, ApJ, 613, 51
Naselsky, P., Zhao, W., Kim, J., & Chen, S. 2012, ApJ, 749, 31
Paci, F., Gruppuso, A., Finelli, F., et al. 2010, MNRAS, 407, 399
Paci, F., Gruppuso, A., Finelli, F., et al. 2013, ArXiv e-prints
Park, C.-G. 2004, MNRAS, 349, 313
Pietrobon, D., Amblard, A., Balbi, A., et al. 2008, Phys. Rev. D, 78, 103504
Planck Collaboration ES. 2013, The Explanatory Supplement to the Planck 2013 results (ESA)
Planck Collaboration I. 2013, Submitted to A&A
Planck Collaboration II. 2013, Submitted to A&A
Planck Collaboration P09A. 2013, A&A
Planck Collaboration VI. 2013, Submitted to A&A
Planck Collaboration XII. 2013, Submitted to A&A
Planck Collaboration XIV. 2013, Submitted to A&A
Planck Collaboration XIX. 2013, Submitted to A&A
Planck Collaboration XV. 2013, Submitted to A&A
Planck Collaboration XVI. 2013, Submitted to A&A
Planck Collaboration XVII. 2013, Submitted to A&A
Planck Collaboration XXIV. 2013, Submitted to A&A
Planck Collaboration XXV. 2013, Submitted to A&A
Planck Collaboration XXVI. 2013, Submitted to A&A
Planck Collaboration XXVII. 2013, Submitted to A&A
Räth, C., Banday, A. J., Rossmanith, G., et al. 2011, MNRAS, 415, 2205
Räth, C., Morfill, G. E., Rossmanith, G., Banday, A. J., & Górski, K. M. 2009, Physical Review Letters, 102, 131301
Räth, C., Schuecker, P., & Banday, A. J. 2007a, MNRAS, 380, 466
Räth, C., Schuecker, P., & Banday, A. J. 2007b, MNRAS, 380, 466
Rossmanith, G., Modest, H., Räth, C., et al. 2012, Phys. Rev. D, 86, 083005
Rossmanith, G., Räth, C., Banday, A. J., & Morfill, G. 2009a, MNRAS, 399, 1921
Rossmanith, G., Räth, C., Banday, A. J., & Morfill, G. 2009b, MNRAS, 399, 1921
Schmalzing, J. & Gorski, K. M. 1998, MNRAS, 297, 355
Schwarz, D. J., Starkman, G. D., Huterer, D., & Copi, C. J. 2004, Physical Review Letters, 93, 221301
Spergel, D. N., Bean, R., Doré, O., et al. 2007, ApJS, 170, 377
Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175
Starobinsky, A. A. 1993, JETP Lett., 57, 622
Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. 2003, Phys.Rev.D, 68, 123523
Vanmarcke, E. 1983, in Random Fields, by Erik Vanmarcke, pp. 372. ISBN 0-262-72045-0. Cambridge, Massachusetts, USA: The MIT Press, March 1983. (Paper), ed. Vanmarcke, E.
Vielva, P. 2010, Advances in Astronomy, 2010
Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, ApJ, 609, 22
Vielva, P., Martínez-González, E., Cruz, M., Barreiro, R. B., & Tucci, M. 2011, MNRAS, 410, 33
Vielva, P., Wiaux, Y., Martínez-González, E., & Vandergheynst, P. 2007, MNRAS, 381, 932
Wiaux, Y., Vielva, P., Barreiro, R. B., Martínez-González, E., & Vandergheynst, P. 2008, MNRAS, 385, 939

Appendix A: Constraints on Alfvén waves

Observations of synchrotron emission and Faraday rotation provide increasing evidence that large-scale astrophysical systems in the Universe are pervaded by magnetic fields. These huge systems include Ly-α forests and intercluster regions (see Kronberg 2009, for a review). Both pre- and post-recombination mechanisms could result in a background of nano-gauss fields that might be detectable in large-scale structures or the CMB, although at present no imprints of these Primordial Magnetic Fields (PMFs) have been detected therein.

Here, we report our findings based on an analysis of the *Planck* data to search for the the predicted signature of statistical anisotropy due to PMFs. Specifically, PMFs with coherence scales comparable to the present day horizon may induce and sustain Alfvén waves in the early Universe that can leave observable imprints on the CMB via the Doppler and integrated Sachs-Wolfe effects. In particular, this results in a preferred angular direction in the CMB anisotropy, therefore breaking statistical isotropy (Durrer et al. 1998, Kim & Naselsky 2009).

Durrer et al. (1998) showed that cosmological Alfvén waves generate a fractional CMB anisotropy for a Fourier mode $\boldsymbol{k}$:

$$\frac{\delta T}{T_0}(\hat{\boldsymbol{n}}, \boldsymbol{k}) \approx \boldsymbol{n} \cdot \boldsymbol{\Omega}(\boldsymbol{k}, \eta_{\rm last}) = \boldsymbol{n} \cdot \boldsymbol{\Omega_0}\; v_A k \eta_{\rm last}\; \boldsymbol{B} \cdot \boldsymbol{k} \tag{A.1}$$

where $\hat{\mathbf{n}}$ denotes sky direction, $\hat{\mathbf{B}}$ is a unit vector in the direction of the coherent PMF, $\boldsymbol{\Omega}(\mathbf{k}, \eta_{\rm last})$ is the Gauge invariant linear combination associated with vector perturbations, $\eta_{\rm last}$ denotes the conformal time at the moment of baryon-photon decoupling, and T_0 is the CMB monopole temperature 2.7255 K (Fixsen 2009). Durrer et al. (1998) assumed that the vector perturbations are initially created by some random stochastic PMF and have the following statistical properties over an ensemble of universes:

$$\langle \Omega_0^i(\mathbf{k})\, \Omega_0^j(\mathbf{k}) \rangle = (\delta_{ij} - \hat{k}_i \hat{k}_j) P(k). \tag{A.2}$$

Here, $P(k)$ is the power spectrum assumed to follow a simple power law:

$$P(k) = A_v \left(\frac{k}{k_0} \right)^{n_v}, \tag{A.3}$$

where k_0 is a pivot wavenumber set to 0.05/Mpc in this analysis. The Alfvén wave velocity is given by (Durrer et al. 1998):

$$v_A = \frac{B}{2\sqrt{\pi(\rho_r + p_r)}} \approx 4 \times 10^{-4} \frac{B}{10^{-9}\text{Gauss}}, \tag{A.4}$$

where ρ_r and p_r are the density and the pressure of photons, and the speed of light c is set to 1.

Kahniashvili et al. (2008) showed that the presence of Alfvén waves in the early Universe leads to specific correlations of the CMB in harmonic space:

$$\begin{aligned}\langle a^*_{\ell m}\, a_{\ell m}\rangle &= C_\ell + \frac{\ell(\ell+1)}{(2\ell-1)(2\ell+3)}\{(\ell^2+\ell-3)\cos^2\theta_B \\ &\quad +\ell(\ell+1) - (3\cos^2\theta_B - 1)m^2\left[1-\frac{3}{\ell(\ell+1)}\right]\} I^{\ell,\ell}_d,\end{aligned}$$

$$\begin{aligned}\langle a^*_{\ell,m\pm1}\, a_{\ell,m}\rangle &= -\sin 2\theta_B \exp[\pm\imath\phi_B]\,\frac{\ell^2+\ell-3}{(2\ell-1)(2\ell+3)}\left(m\pm\frac{1}{2}\right) \\ &\quad\times\ \sqrt{(\ell\mp m)(\ell\pm m+1)}\, I^{\ell,\ell}_d,\end{aligned}$$

$$\begin{aligned}\langle a^*_{\ell,m\pm2}\, a_{\ell,m}\rangle &= -\frac{1}{2}\sin^2\theta_B \exp[\pm\imath 2\phi_B]\,\frac{\ell^2+\ell-3}{(2\ell-1)(2\ell+3)} \\ &\quad\times\ \sqrt{(\ell\mp m)(\ell\mp m-1)(\ell\pm m+1)(\ell\pm m+2)}\, I^{\ell,\ell}_d,\end{aligned}$$

$$\begin{aligned}\langle a^*_{\ell+2,m}\, a_{\ell,m}\rangle &= -(3\cos^2\theta_B-1)\frac{(\ell+3)\ell}{2(2\ell+3)\sqrt{(2\ell+1)(2\ell+5)}} \\ &\quad\times\ \sqrt{((\ell+1)^2-m^2)((\ell+2)^2-m^2)}\, I^{\ell,\ell+2}_d,\end{aligned}$$

$$\begin{aligned}\langle a^*_{\ell-2,m}\, a_{\ell,m}\rangle &= -(3\cos^2\theta_B-1)\frac{(\ell+1)(\ell-2)}{2(2\ell-1)\sqrt{(2\ell-3)(2\ell+1)}} \\ &\quad\times\ \sqrt{((\ell-1)^2-m^2)(\ell^2-m^2)}\, I^{\ell,\ell-2}_d,\end{aligned}$$

$$\begin{aligned}\langle a^*_{\ell+2,m\pm1}\, a_{\ell,m}\rangle &= \sin 2\theta_B \exp[\pm\imath\phi_B]\frac{(\ell+3)l}{2(2\ell+3)\sqrt{(2\ell+1)(2\ell+5)}} \\ &\quad\times\ \sqrt{((\ell+1)^2-m^2)(\ell\pm m+2)(\ell\pm m+3)}\, I^{\ell,\ell+2}_d,\end{aligned}$$

$$\begin{aligned}&\langle a^*_{\ell-2,m\pm1}\, a_{\ell,m}\rangle = \\ &\quad -\sin 2\theta_B \exp[\pm\imath\phi_B]\frac{(\ell+1)(\ell-2)}{2(2\ell-1)\sqrt{(2\ell-3)(2\ell+1)}} \\ &\quad\times\sqrt{(\ell^2-m^2)(\ell\mp m-1)(\ell\mp m-2)}\, I^{\ell,\ell-2}_d,\end{aligned}$$

$$\begin{aligned}&\langle a^*_{\ell+2,m\pm2}\, a_{\ell,m}\rangle = \\ &\quad -\frac{1}{2}\sin^2\theta_B \exp[\pm\imath 2\phi_B]\frac{(\ell+3)\ell}{2(2\ell+3)\sqrt{(2\ell+1)(2\ell+5)}} \\ &\times\ \sqrt{((\ell\pm m+1)(\ell\pm m+2)(\ell\pm m+3)(\ell\pm m+4)}\, I^{\ell,\ell+2}_d,\end{aligned}$$

$$\begin{aligned}&\langle a^*_{\ell-2,m\pm2}\, a_{\ell,m}\rangle = \\ &\quad -\frac{1}{2}\sin^2\theta_B \exp[\pm\imath 2\phi_B]\frac{(\ell+1)(\ell-2)}{2(2\ell-1)\sqrt{(2\ell-3)(2\ell+1)}} \\ &\quad\times\sqrt{((\ell\mp m-3)(\ell\mp m-2)(\ell\mp m-1)(\ell\mp m)}\, I^{\ell,\ell-2}_d,\end{aligned}$$

where C_ℓ is the power spectrum in the absence of Alfvén waves, θ_B and ϕ_B are the spherical angles of a PMF direction $\boldsymbol{B}$. Here, $I^{\ell\ell'}_d$ is given by:

$$\begin{aligned}I^{\ell\ell'}_d &= \frac{2\,T_0^2}{\pi}\int d\ln k\, k^3 A_v\left(\frac{k}{k_0}\right)^{n_v} \\ &\quad\times\ \exp\left(-2\frac{k^2}{k_D^2}\right) v_A^2\left(\frac{\eta_{\rm last}}{\eta_0}\right)^2 j_\ell(k\eta_0)\, j_{\ell'}(k\eta_0) \\ &= \frac{2T_0^2}{\pi}\left(\frac{\eta_{\rm last}}{\eta_0}\right)^2 A_v v_A^2\, k_0^3 \\ &\quad\times\ \int d\ln k\left(\frac{k}{k_0}\right)^{n_v+3}\exp\left(-2\frac{k^2}{k_D^2}\right) j_\ell(k\eta_0)\, j_{\ell'}(k\eta_0),\end{aligned}$$

where η_0 is the present conformal time, and k_D denotes the comoving wavenumber of the dissipation scale, due to photon viscosity and given by approximately $10/\eta_{\rm last}$ (Durrer et al. 1998). The damping effect becomes significant on multipoles $\ell \gtrsim 500$ (Durrer et al. 1998). As shown above, Alfvén waves in the early Universe produce correlations between harmonic modes separated by $\Delta\ell = 0, \pm2$, and $\Delta m = 0, \pm2$. Investigating these imprints, we may impose a constraint on the Alfvén waves. In the weak Alfvén wave limit, the CMB data likelihood can be expanded as follows:

$$\begin{aligned}\mathcal{L} \approx\ & \mathcal{L}|_{A_v v_A^2=0} + \left.\frac{\partial\mathcal{L}}{\partial(A_v v_A^2)}\right|_{A_v v_A^2=0} A_v v_A^2 + \frac{1}{2}\left.\frac{\partial^2\mathcal{L}}{\partial(A_v v_A^2)^2}\right|_{A_v v_A^2=0}(A_v v_A^2)^2 \\ & +O((A_v v_A^2)^3).\end{aligned} \tag{A.5}$$

Since all correlations produced by Alfvén waves are proportional to $A_v v_A^2$, the first term in Eq. A.5 is simply equal to the likelihood of the standard cosmological model. The first and second derivative of the likelihood are obtained by:

$$\frac{\partial\mathcal{L}}{\partial\lambda} = \mathcal{H} - \langle\mathcal{H}\rangle, \qquad \frac{\partial^2\mathcal{L}}{\partial\lambda^2} = -\langle\mathcal{H}^2\rangle + \langle\mathcal{H}\rangle\langle\mathcal{H}\rangle, \tag{A.6}$$

where

$$\mathcal{H} = \frac{1}{2}\left[\mathsf{C}^{-1}\boldsymbol{a}\right]^\dagger \frac{\partial\mathsf{C}}{\partial\lambda}\left[\mathsf{C}^{-1}\boldsymbol{a}\right], \tag{A.7}$$

$\boldsymbol{a}$ is the data vector, consisting of the spherical harmonic coefficients, $a_{\ell m}$, of the CMB anisotropy data, and $\mathbf{C}$ is their covariance matrix.

In our analysis, we consider the four foreground cleaned CMB maps `Commander-Ruler`, `NILC`, `SEVEM`, and `SMICA` and apply the common mask. We assume the fiducial *Planck* cosmological model, and use MC simulations to estimate the ensemble average values for signal and noise, as required in Eq. A.6. The quantity $\mathsf{C}^{-1}\boldsymbol{a}$ from Eq. A.7 was then determined by the messenger field method (Elsner & Wandelt 2013). The `CosmoMC` package (Lewis & Bridle 2002) is then used as a generic sampler in order to obtain the posterior probability for the Alfvén wave parameters $\{A_v v_A^2, n_v, \theta_B, \phi_B\}$.

In Table A.1, we show the upper bounds on the Alfvén wave amplitude $A_v v_A^2$ at various confidence levels, after marginalizing over the spectral index n_v and the direction θ_B, ϕ_B. From the analysis of the *Planck* data, we impose an upper bound on the Alfvén wave amplitude that is tighter than that from the *WMAP* data by more than one order of magnitude.

[1] APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/lrfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex

Table A.1. *Planck* constraints on the Alfvén wave amplitude $A_v v_A^2$.

Confidence Level	68%	95%	99.7%
C-R	$< 0.48 \times 10^{-9}$	$< 1.01 \times 10^{-9}$	$< 1.57 \times 10^{-9}$
NILC	$< 0.49 \times 10^{-9}$	$< 1.00 \times 10^{-9}$	$< 1.56 \times 10^{-9}$
SEVEM	$< 0.54 \times 10^{-9}$	$< 1.13 \times 10^{-9}$	$< 1.73 \times 10^{-9}$
SMICA	$< 0.47 \times 10^{-9}$	$< 0.87 \times 10^{-9}$	$< 1.29 \times 10^{-9}$

13, France

[2] Aalto University Metsähovi Radio Observatory, Metsähovintie 114, FIN-02540 Kylmälä, Finland

[3] African Institute for Mathematical Sciences, 6-8 Melrose Road, Muizenberg, Cape Town, South Africa

[4] Agenzia Spaziale Italiana Science Data Center, c/o ESRIN, via Galileo Galilei, Frascati, Italy

[5] Agenzia Spaziale Italiana, Viale Liegi 26, Roma, Italy

[6] Astrophysics Group, Cavendish Laboratory, University of Cambridge, J J Thomson Avenue, Cambridge CB3 0HE, U.K.

[7] Astrophysics & Cosmology Research Unit, School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Westville Campus, Private Bag X54001, Durban 4000, South Africa

[8] CITA, University of Toronto, 60 St. George St., Toronto, ON M5S 3H8, Canada

[9] CNRS, IRAP, 9 Av. colonel Roche, BP 44346, F-31028 Toulouse cedex 4, France

[10] California Institute of Technology, Pasadena, California, U.S.A.

[11] Centre for Theoretical Cosmology, DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA U.K.

[12] Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Plaza San Juan, 1, planta 2, E-44001, Teruel, Spain

[13] Computational Cosmology Center, Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.

[14] Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain

[15] DSM/Irfu/SPP, CEA-Saclay, F-91191 Gif-sur-Yvette Cedex, France

[16] DTU Space, National Space Institute, Technical University of Denmark, Elektrovej 327, DK-2800 Kgs. Lyngby, Denmark

[17] Département de Physique Théorique, Université de Genève, 24, Quai E. Ansermet,1211 Genève 4, Switzerland

[18] Departamento de Física Fundamental, Facultad de Ciencias, Universidad de Salamanca, 37008 Salamanca, Spain

[19] Departamento de Física, Universidad de Oviedo, Avda. Calvo Sotelo s/n, Oviedo, Spain

[20] Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Avda. de los Castros s/n, Santander, Spain

[21] Department of Astronomy and Astrophysics, University of Toronto, 50 Saint George Street, Toronto, Ontario, Canada

[22] Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

[23] Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California, U.S.A.

[24] Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia, Canada

[25] Department of Physics and Astronomy, Dana and David Dornsife College of Letter, Arts and Sciences, University of Southern

California, Los Angeles, CA 90089, U.S.A.

[26] Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.

[27] Department of Physics, Gustaf Hällströmin katu 2a, University of Helsinki, Helsinki, Finland

[28] Department of Physics, Princeton University, Princeton, New Jersey, U.S.A.

[29] Department of Physics, University of Alberta, 11322-89 Avenue, Edmonton, Alberta, T6G 2G7, Canada

[30] Department of Physics, University of California, Berkeley, California, U.S.A.

[31] Department of Physics, University of California, One Shields Avenue, Davis, California, U.S.A.

[32] Department of Physics, University of California, Santa Barbara, California, U.S.A.

[33] Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois, U.S.A.

[34] Dipartimento di Fisica e Astronomia G. Galilei, Università degli Studi di Padova, via Marzolo 8, 35131 Padova, Italy

[35] Dipartimento di Fisica e Scienze della Terra, Università di Ferrara, Via Saragat 1, 44122 Ferrara, Italy

[36] Dipartimento di Fisica, Università La Sapienza, P. le A. Moro 2, Roma, Italy

[37] Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria, 16, Milano, Italy

[38] Dipartimento di Fisica, Università degli Studi di Trieste, via A. Valerio 2, Trieste, Italy

[39] Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy

[40] Dipartimento di Matematica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, 1, Roma, Italy

[41] Discovery Center, Niels Bohr Institute, Blegdamsvej 17, Copenhagen, Denmark

[42] Dpto. Astrofísica, Universidad de La Laguna (ULL), E-38206 La Laguna, Tenerife, Spain

[43] European Space Agency, ESAC, Planck Science Office, Camino bajo del Castillo, s/n, Urbanización Villafranca del Castillo, Villanueva de la Cañada, Madrid, Spain

[44] European Space Agency, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands

[45] Helsinki Institute of Physics, Gustaf Hällströmin katu 2, University of Helsinki, Helsinki, Finland

[46] INAF - Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, Padova, Italy

[47] INAF - Osservatorio Astronomico di Roma, via di Frascati 33, Monte Porzio Catone, Italy

[48] INAF - Osservatorio Astronomico di Trieste, Via G.B. Tiepolo 11, Trieste, Italy

[49] INAF/IASF Bologna, Via Gobetti 101, Bologna, Italy

[50] INAF/IASF Milano, Via E. Bassini 15, Milano, Italy

[51] INFN, Sezione di Bologna, Via Irnerio 46, I-40126, Bologna, Italy

[52] INFN, Sezione di Roma 1, Università di Roma Sapienza, Piazzale Aldo Moro 2, 00185, Roma, Italy

[53] IPAG: Institut de Planétologie et d'Astrophysique de Grenoble, Université Joseph Fourier, Grenoble 1 / CNRS-INSU, UMR 5274,

Grenoble, F-38041, France

[54] ISDC Data Centre for Astrophysics, University of Geneva, ch. d'Ecogia 16, Versoix, Switzerland

[55] IUCAA, Post Bag 4, Ganeshkhind, Pune University Campus, Pune 411 007, India

[56] Imperial College London, Astrophysics group, Blackett Laboratory, Prince Consort Road, London, SW7 2AZ, U.K.

[57] Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, U.S.A.

[58] Institut Néel, CNRS, Université Joseph Fourier Grenoble I, 25 rue des Martyrs, Grenoble, France

[59] Institut Universitaire de France, 103, bd Saint-Michel, 75005, Paris, France

[60] Institut d'Astrophysique Spatiale, CNRS (UMR8617) Université Paris-Sud 11, Bâtiment 121, Orsay, France

[61] Institut d'Astrophysique de Paris, CNRS (UMR7095), 98 bis Boulevard Arago, F-75014, Paris, France

[62] Institute for Space Sciences, Bucharest-Magurale, Romania

[63] Institute of Astronomy and Astrophysics, Academia Sinica, Taipei, Taiwan

[64] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, U.K.

[65] Institute of Theoretical Astrophysics, University of Oslo, Blindern, Oslo, Norway

[66] Instituto de Astrofísica de Canarias, C/Vía Láctea s/n, La Laguna, Tenerife, Spain

[67] Instituto de Física de Cantabria (CSIC-Universidad de Cantabria), Avda. de los Castros s/n, Santander, Spain

[68] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California, U.S.A.

[69] Jodrell Bank Centre for Astrophysics, Alan Turing Building, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester, M13 9PL, U.K.

[70] Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 0HA, U.K.

[71] LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France

[72] LERMA, CNRS, Observatoire de Paris, 61 Avenue de l'Observatoire, Paris, France

[73] Laboratoire AIM, IRFU/Service d'Astrophysique - CEA/DSM - CNRS - Université Paris Diderot, Bât. 709, CEA-Saclay, F-91191 Gif-sur-Yvette Cedex, France

[74] Laboratoire Traitement et Communication de l'Information, CNRS (UMR 5141) and Télécom ParisTech, 46 rue Barrault F-75634 Paris Cedex 13, France

[75] Laboratoire de Physique Subatomique et de Cosmologie, Université Joseph Fourier Grenoble I, CNRS/IN2P3, Institut National Polytechnique de Grenoble, 53 rue des Martyrs, 38026 Grenoble cedex, France

[76] Laboratoire de Physique Théorique, Université Paris-Sud 11 & CNRS, Bâtiment 210, 91405 Orsay, France

[77] Lawrence Berkeley National Laboratory, Berkeley, California, U.S.A.

[78] Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany

[79] Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstraße, 85748 Garching, Germany

[80] McGill Physics, Ernest Rutherford Physics Building, McGill University, 3600 rue University, Montréal, QC, H3A 2T8, Canada

[81] MilliLab, VTT Technical Research Centre of Finland, Tietotie 3, Espoo, Finland

[82] Niels Bohr Institute, Blegdamsvej 17, Copenhagen, Denmark

[83] Observational Cosmology, Mail Stop 367-17, California Institute of Technology, Pasadena, CA, 91125, U.S.A.

[84] Optical Science Laboratory, University College London, Gower Street, London, U.K.

[85] SB-ITP-LPPC, EPFL, CH-1015, Lausanne, Switzerland

[86] SISSA, Astrophysics Sector, via Bonomea 265, 34136, Trieste, Italy

[87] School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff, CF24 3AA, U.K.

[88] School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, U.K.

[89] Space Sciences Laboratory, University of California, Berkeley, California, U.S.A.

[90] Special Astrophysical Observatory, Russian Academy of Sciences, Nizhnij Arkhyz, Zelenchukskiy region, Karachai-Cherkessian Republic, 369167, Russia

[91] Stanford University, Dept of Physics, Varian Physics Bldg, 382 Via Pueblo Mall, Stanford, California, U.S.A.

[92] Sub-Department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, U.K.

[93] Theory Division, PH-TH, CERN, CH-1211, Geneva 23, Switzerland

[94] UPMC Univ Paris 06, UMR7095, 98 bis Boulevard Arago, F-75014, Paris, France

[95] Université de Toulouse, UPS-OMP, IRAP, F-31028 Toulouse cedex 4, France

[96] University of Granada, Departamento de Física Teórica y del Cosmos, Facultad de Ciencias, Granada, Spain

[97] University of Miami, Knight Physics Building, 1320 Campo Sano Dr., Coral Gables, Florida, U.S.A.

[98] Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland

Paper IV

CMB likelihood approximation for banded probability distributions

Gjerløw, E. and Mikkelsen, K. and Eriksen, H. K. and Górski, K. M. and Huey, G. and Jewell, J. B. and Næss, S. K. and Rocha, G. and Seljebotn, D. S. and Wehus, I. K.

Astrophysical Journal, Volume 777, issue 2, article id. 150 (2013)

The Astrophysical Journal, 777:150 (9pp), 2013 November 10 doi:10.1088/0004-637X/777/2/150

COSMIC MICROWAVE BACKGROUND LIKELIHOOD APPROXIMATION FOR BANDED PROBABILITY DISTRIBUTIONS

E. Gjerløw[1], K. Mikkelsen[1], H. K. Eriksen[1], K. M. Górski[2,3], G. Huey[2], J. B. Jewell[2], S. K. Næss[1,4], G. Rocha[2,5], D. S. Seljebotn[1], and I. K. Wehus[2]
[1] Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, N-0315 Oslo, Norway; eirik.gjerlow@astro.uio.no
[2] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[3] Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland
[4] Department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, UK
[5] California Institute of Technology, Pasadena, CA, USA
Received 2013 March 30; accepted 2013 September 17; published 2013 October 23

ABSTRACT

We investigate sets of random variables that can be arranged sequentially such that a given variable only depends conditionally on its immediate predecessor. For such sets, we show that the full joint probability distribution may be expressed exclusively in terms of uni- and bivariate marginals. Under the assumption that the cosmic microwave background (CMB) power spectrum likelihood only exhibits correlations within a banded multipole range, $\Delta\ell_C$, we apply this expression to two outstanding problems in CMB likelihood analysis. First, we derive a statistically well-defined hybrid likelihood estimator, merging two independent (e.g., low- and high-ℓ) likelihoods into a single expression that properly accounts for correlations between the two. Applying this expression to the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) likelihood, we verify that the effect of correlations on cosmological parameters in the transition region is negligible in terms of cosmological parameters for *WMAP*; the largest relative shift seen for any parameter is 0.06σ. However, because this may not hold for other experimental setups (e.g., for different instrumental noise properties or analysis masks), but must rather be verified on a case-by-case basis, we recommend our new hybridization scheme for future experiments for statistical self-consistency reasons. Second, we use the same expression to improve the convergence rate of the Blackwell–Rao likelihood estimator, reducing the required number of Monte Carlo samples by several orders of magnitude, and thereby extend it to high-ℓ applications.

Key words: cosmic background radiation – cosmology: observations – methods: data analysis – methods: numerical – methods: statistical

Online-only material: color figures

1. INTRODUCTION

The cosmic microwave background (CMB) radiation is one of the most pristine sources of information about the early universe available to us. Since its discovery in 1964 (Penzias & Wilson 1965), the amount of information available to us about the CMB has increased at a rapid pace through series of ground-based, sub-orbital and satellite experiments. The recently released *Planck* temperature sky maps (Planck I 2013) is just the latest example of how the present challenge in the field of cosmology is one of overabundance rather than shortage of data.

To extract cosmological parameters from these ever growing data sets requires increasingly sophisticated and efficient algorithms, both due to larger data volumes and to more stringent requirements to statistical precision. For example, the *COBE*-DMR sky maps published 20 yr ago (Smoot et al. 1992) comprised $\mathcal{O}(10^4)$ pixels, and could be analyzed using exact brute-force likelihood techniques (e.g., Górski 1994), with a computational scaling of $\mathcal{O}(N_{\rm pix}^3)$. The *Wilkinson Microwave Anisotropy Probe* (*WMAP*) sky maps published 10 yr ago comprised $\mathcal{O}(10^7)$ pixels (Bennett et al. 2003a), at which point faster and approximate methods had to be used for parameter estimation (Hivon et al. 2002; Verde et al. 2003). However, for *WMAP* the error budget was still dominated by cosmic variance on large angular scales and instrumental noise on small angular scales, and confusion with Galactic and extra-Galactic emission was minimal, allowing for very simple component separation methods (Bennett et al. 2003b; Hinshaw et al. 2003). For *Planck*, the total number of data points in nine frequency bands is $\mathcal{O}(3\times10^8)$, and instrumental noise never dominates the uncertainties at any angular scales, as small-scale astrophysical confusion becomes important at multipoles $\ell\gtrsim1500$ (Planck XII 2013). As a result, an unprecedented study of all important sources of uncertainty, including instrumental, systematic and astrophysical, was required for *Planck* to reach its ambitious goals (Planck XV 2013).

With the advent of these massive mega-pixel data sets, a number different analysis strategies have been developed to robustly extract cosmological parameters with acceptable computational cost. As of today, the preferred option for full-sky high-resolution experiments such as *Planck* and *WMAP* is to divide the analysis into two separate components according to large and small angular scales, and merge the two at the likelihood level. On large angular scales, they use a Gibbs sampling based (Jewell et al. 2004; Wandelt et al. 2004; Eriksen et al. 2004) Blackwell–Rao (BR) estimator (Chu et al. 2005) that takes into account the full non-Gaussian structure of the true CMB likelihood, while on small angular scales, they use faster approaches (e.g., Hivon et al. 2002; Rocha et al. 2011; Planck XV 2013) coupled to an analytic multivariate Gaussian (and/or log-normal) likelihood approximation. The computational cost of this hybrid approach is dominated by spherical harmonics transforms, and therefore scales as $\mathcal{O}(N_{\rm pix}^{3/2})$, which is acceptable even for large data sets. However, there is an unsolved problem

associated with this hybrid approach, and that is how to merge the two likelihood components into a single all-scale expression; correlations between the smallest scales in the large-scale likelihood and the largest scales in the small-scale likelihood should in principle be accounted for. As of today, no fully satisfactory solution to this exists in the CMB literature, though various approaches were explored during the Planck analysis.

Having a computational scaling of $\mathcal{O}(N_{\rm pix}^{3/2})$, the Gibbs sampling approach could in principle be employed for all angular scales, thus eliminating the need for any hybrid approximation. Unfortunately, in practice this method is in its current implementation limited to low angular scales for two reasons: First, joint CMB analysis and component separation is currently implemented in terms of pixel-based fits of physical foreground models, requiring all frequency bands to have the same angular resolution, dictated by the coarsest resolution in a given data set. Second, although the computational scaling for the Gibbs sampler is acceptable, the prefactor is high. The 2013 *Planck* likelihood employed 100,000 Gibbs samples in order to achieve robust BR convergence, and each of those samples required ∼2000 Conjugate Gradient iterations (and twice as many spherical harmonic transforms) to converge, for a total cost of 500,000 CPU hours. Naively scaling this to full *Planck* resolution suggest a final cost of $\mathcal{O}(10^8)$ CPU hours (only taking into account the additional computational cost per sample for high-resolution Gibbs sampling, not the additional number of samples needed for the BR estimator to converge at higher multipoles).

The main result of the present paper is a statistically well motivated block factorization of the CMB power spectrum likelihood that is applicable to several of these problems. Specifically, we show that for sets of random variables that can be arranged sequentially in such a way that all correlations have a finite range within the sequence, the full joint probability distribution may be written in terms of lower-dimensional marginals. The archetypal example of such a distribution is a multivariate Gaussian with a strictly banded covariance matrix, and we therefore call the general (non-Gaussian but conditionally limited) case also "banded." With this statistical identity ready at hand, we first suggest a statistically well-motivated likelihood hybridization scheme that takes properly into account correlations between the low- and high-ℓ regimes, and, second, we show how the convergence rate of the BR estimator can be improved by factorizing the full high-dimensional multivariate posterior into a set of lower-dimensional distributions, each of which converges much faster than the full distribution. This approach differs from the direct Gaussianization technique proposed by Rudjord et al. (2009) in that the underlying probabilistic structure (e.g., shapes of marginal and N-point correlations) is conserved; in principle, the only modification to the full likelihood enforced by our new approach is that assumed negligible correlations are explicitly set to zero.

2. FACTORIZING THE CMB LIKELIHOOD

2.1. Factorization of Banded Probability Distributions

We begin with a general joint probability density $P(\{\theta\}) = P(\theta_1, \theta_2, \theta_3, \ldots, \theta_n)$ for a set of random variables, θ_k, with $k = 1, 2, 3, \ldots, n$. We choose one specific sequential ordering of these variables (out of all the possible orderings), and use the definition of a conditional to write the joint distribution as a product of univariate conditionals,

$$\begin{aligned} P(\{\theta\}) = &\ P(\theta_1, \theta_2, \theta_3, \ldots, \theta_n) \\ = &\ P(\theta_1|\theta_2, \theta_3, \ldots, \theta_n) \\ &\cdot P(\theta_2|\theta_3, \ldots \theta_n) \cdots \\ &\cdot P(\theta_{n-1}|\theta_n) \cdot P(\theta_n). \end{aligned}$$

We then assume that our variables only have a conditional probability dependence on their immediate neighbors in the sequence, i.e., that the probability distribution is *tri-diagonally* banded,

$$\begin{aligned} P(\{\theta\}) &\approx P(\theta_1|\theta_2) \cdot P(\theta_2|\theta_3) \cdots P(\theta_{n-1}|\theta_n) \cdot P(\theta_n) \\ &= \frac{P(\theta_1, \theta_2)}{P(\theta_2)} \cdot \frac{P(\theta_2, \theta_3)}{P(\theta_3)} \cdots \frac{P(\theta_{n-1}, \theta_n)}{P(\theta_n)} \cdot P(\theta_n) \\ &= \frac{\prod_{k=1}^{n-1} P(\theta_k, \theta_{k+1})}{\prod_{k=2}^{n-1} P(\theta_k)}. \end{aligned} \quad (1)$$

Thus, this simple derivation shows that a strictly (tri-diagonally) banded probability distribution may be factorized recursively into a product of uni- and bivariate marginals.

Before applying this expression to CMB likelihood approximation, we note that even if the joint probability distribution do not have correlations exclusively between neighboring variables, it may still be possible to factorize it, provided at least some correlations may be ignored. For instance, suppose we can ignore all but the nearest *two* neighbors; in that case, the joint distribution will factorize into a product of uni-, bi- and trivariate marginals.

2.2. Block Factorization of the CMB Likelihood

In its most basic representation, a CMB data set, $\mathbf{d}$, may be modeled as

$$\mathbf{d} = \mathbf{s} + \mathbf{n}, \quad (2)$$

where $\mathbf{s}$ is the true sky signal and $\mathbf{n}$ represents instrumental noise. Both the signal and noise are usually assumed to be zero-mean Gaussian variables with covariances $\mathbf{S}$ and $\mathbf{N}$, respectively.

The noise covariance matrix is typically given by external knowledge about the instrumental noise characteristics and the scanning strategy of a given experiment. The signal covariance matrix, on the other hand, is generally unknown, and must be estimated from the data. However, given the fact that we only have one observable sky available, it is impossible to estimate the $N_{\rm pix}^2$ elements in $\mathbf{S}$ from the $N_{\rm pix}$ elements in $\mathbf{d}$ without imposing strong priors on its structure. The most commonly accepted prior is simply that the CMB sky is isotropic and homogeneous (e.g., Planck XXIII 2013). It is therefore convenient to expand $\mathbf{s}$ in spherical harmonics, such that

$$\mathbf{s}(\hat{\mathbf{n}}) = \sum_{\ell,m} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \quad (3)$$

where $\hat{\mathbf{n}}$ is a unit vector pointing to a given position on the sky, $Y_{\ell m}$ are the spherical harmonics, and $a_{\ell m}$ are the corresponding spherical harmonics coefficients. Then the signal covariance matrix may be written as

$$S_{\ell m, \ell' m'} = \langle a_{\ell m} a^*_{\ell' m'} \rangle \equiv C_\ell \delta_{\ell \ell'} \delta_{m m'}, \quad (4)$$

where C_ℓ is known as the angular power spectrum.

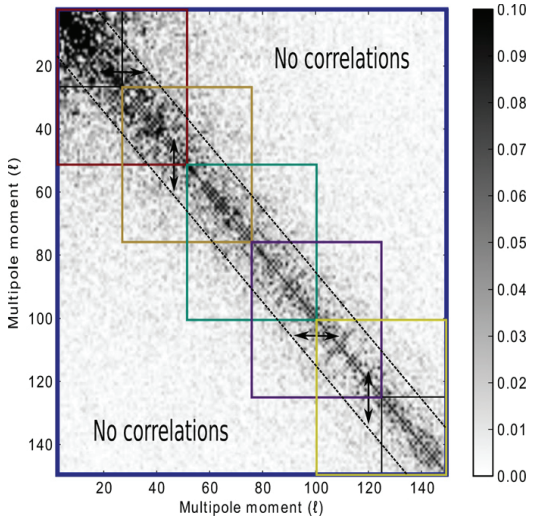


Figure 1. Angular power spectrum correlation matrix, $M_{\ell\ell'}$, for the official *Planck* low-ℓ CMB data set, estimated by Monte Carlo sampling. Note that any two-point correlations are contained within a band of $\Delta\ell_C \sim 15$, suggesting that the CMB likelihood may be approximated as a banded probability distribution. To factorize the CMB likelihood into lower-dimensional elements, we partition the full multipole range into a set of disjoint blocks such that all non-zero covariance elements are embedded within a tri-diagonal block structure, indicated here by colored squares.

(A color version of this figure is available in the online journal.)

The main goal of most CMB experiments is precisely to measure the CMB power spectrum, and the most straightforward way to do so is by maximum-likelihood estimation. Since we have assumed that both signal and noise are Gaussian distributed, the CMB power spectrum likelihood simply reads

$$\mathcal{L}(C_\ell) \equiv P(\mathbf{d}|C_\ell) \propto \frac{e^{-\frac{1}{2}\mathbf{d}^t(\mathbf{S}(C_\ell)+\mathbf{N})^{-1}\mathbf{d}}}{\sqrt{|\mathbf{S}(C_\ell)+\mathbf{N}|}}, \quad (5)$$

where $\mathbf{S} = \mathbf{S}(C_\ell)$ is the covariance matrix given in Equation (4) expressed in pixel domain. Note that C_ℓ denotes the set of all power spectrum coefficients, and the likelihood therefore spans an $\ell_{\rm max}$-dimensional space.

As already mentioned, brute-force evaluation of Equation (5) scales computationally as $\mathcal{O}(N_{\rm pix}^3)$, and is therefore feasible only for very low angular resolutions. Much of the CMB analysis literature therefore revolves around finding computationally tractable approximations to this expression.

In order to build up some intuition about the correlation structure of $\mathcal{L}(C_\ell)$, it is useful to plot the correlation matrix

$$M_{\ell\ell'} \equiv \frac{\langle (C_\ell - \langle C_\ell\rangle)(C_{\ell'} - \langle C_{\ell'}\rangle)\rangle}{\sqrt{\langle (C_\ell - \langle C_\ell\rangle)^2\rangle \langle (C_{\ell'} - \langle C_{\ell'}\rangle)^2\rangle}}. \quad (6)$$

Figure 1 shows this matrix for the official *Planck* low-ℓ CMB data, as evaluated from 200,000 Monte Carlo samples generated with a CMB Gibbs sampler (Jewell et al. 2004; Wandelt et al. 2004; Eriksen et al. 2004). In this case, there are significant correlations between all elements at $\ell \lesssim 20$, while at $\ell \gtrsim 50$ any correlations are well contained inside a band of $\Delta\ell_C = 15$; any correlations beyond $\Delta\ell_C \gtrsim 30$ are well below 1%. Higher-order correlations are significantly smaller than these two-point correlations.

For typical sky cuts and instrumental noise characteristics, the basic CMB likelihood can therefore be approximated as a banded probability distribution with a bandwidth of $\ell \lesssim 15$, and can therefore in principle be factorized by Equation (1). However, as currently written this expression only applies to a strictly tri-diagonal covariance matrix. To circumvent this problem, we therefore introduce an auxiliary block structure that embeds all non-negligible elements within a larger tri-diagonal structure, as illustrated by the colored blocks in Figure 1. That is, we define a set of multipole blocks such that $\theta_1 = \{C_{\ell_{\rm min}}, \ldots, C_{\ell_1}\}$, $\theta_2 = \{C_{\ell_1+1}, \ldots, C_{\ell_2}\}$,..., $\theta_n = \{C_{\ell_{n-1}+1}, \ldots, C_{\ell_{\rm max}}\}$. Thus, each univariate marginal in Equation (1) is replaced with a multivariate distribution of dimension $\ell_i - \ell_{i-1}$, and each bivariate marginal is replaced with a multivariate distribution of dimension $\ell_i - \ell_{i-2}$. This block-wise factorization constitutes the main result of this paper, and in the following sections we will apply this to two concrete problems in CMB likelihood estimation.

3. ACCURATE HYBRID CMB LIKELIHOOD ESTIMATION

As already mentioned, both *Planck* and *WMAP* have adopted so-called "hybrid" likelihood approximations, combining a Gibbs sampling based BR estimator at large angular scales with a Gaussian (and/or log-normal) pseudo cross-spectrum approximation at small angular scales. These two components are merged into a single expression at the log-likelihood level. The *Planck* likelihood simply adds the two log-likelihoods (Planck XV 2013), adopting a so-called "sharp transition" between the low- and high-ℓ regimes, schematically illustrated in the left panel of Figure 2. This is the simplest possible approach, and assumes that any correlations across the transition multipole are negligible. The *WMAP* likelihood makes a different choice, by including the off-diagonal terms between the low- and high-ℓ blocks in the (Gaussian plus log-normal) high-ℓ likelihood, as illustrated in the middle panel of Figure 2.

In this section, we introduce a new and statistically better motivated approach than either of two employed by *Planck* and *WMAP*, taking advantage of the block factorization derived in Equation (1). The first step in our approach is to partition the full multipole range between $\ell_{\rm min}$ and $\ell_{\rm min}$ into three disjoint regions, $L = \{\ell_{\rm min}, \ldots, \ell_{\rm low}\}$, $T = \{\ell_{\rm low}+1, \ldots, \ell_{\rm high}-1\}$ and $H = \{\ell_{\rm high}, \ldots, \ell_{\rm max}\}$, corresponding to a low-ℓ region, a transition region and a high-ℓ region, respectively. The width of the transition region is chosen to be at least as wide as the effective bandwidth of the C_ℓ covariance matrix (see Figure 1). With this partitioning, we now specialize Equation (1) to the case with $n = 3$ regions:

$$\log \mathcal{L}(C_\ell) = \log \mathcal{L}(L, T) + \log \mathcal{L}(T, H) - \log \mathcal{L}(T). \quad (7)$$

Note that this approximation is exact under the assumption of vanishing correlations between the low- and high-ℓ regions, which can be ensured simply by letting the transition region be sufficiently wide. This estimator is schematically illustrated in the right panel of Figure 2.

Equation (7) has a simple intuitive interpretation: The log-likelihood is simply the sum of a low- and a high-ℓ contribution, defined such that they overlap over a sufficiently wide multipole range that all non-negligible correlations are included. However, because the diagonal block in the transition region

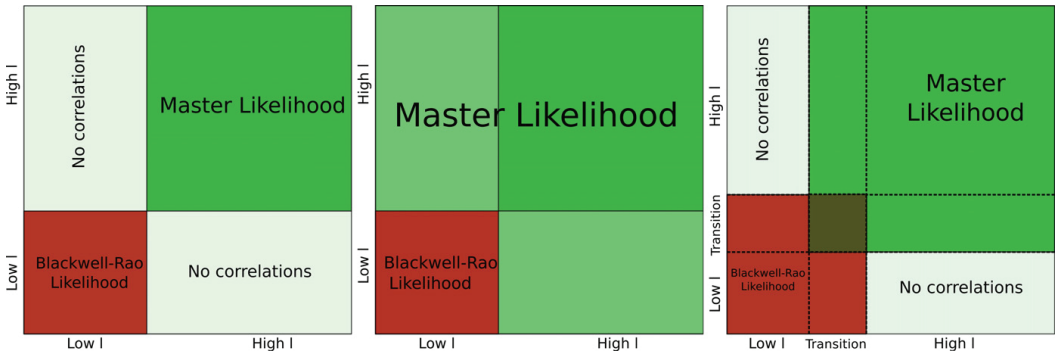


Figure 2. Schematic overview of the three hybridization schemes discussed in the text. The left panel illustrates a sharp transition between the low- (Blackwell–Rao) and high-ℓ (MASTER) likelihood, as currently adopted by *Planck*. The middle panel illustrates the *WMAP* approach, which includes the off-diagonal elements between the low- and high-ℓ regions in the high-ℓ likelihood estimator. The right panel illustrates the new estimator proposed in this paper, in which correlations are accounted for through an transition region that is sufficiently wide to include all non-negligible correlations between the low- and high-ℓ regions. To avoid double-counting of the diagonal elements, the total log-likelihood is corrected by the log-likelihood including elements within the transition region only.

(A color version of this figure is available in the online journal.)

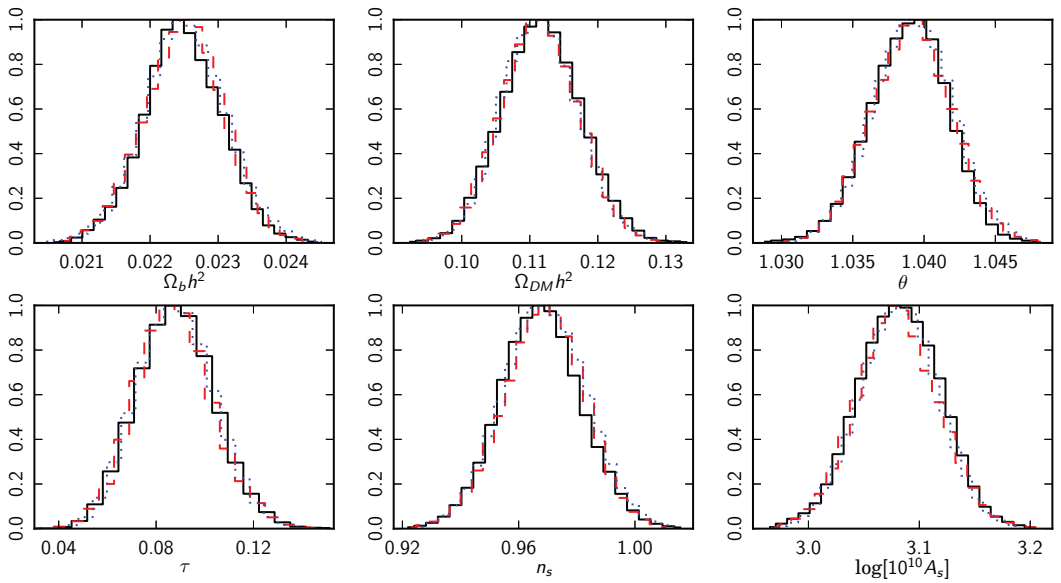


Figure 3. Comparison of best-fit parameters derived by CosmoMC from *WMAP* using likelihood approximations based on the new hybrid estimator presented in this paper (solid black line); the *WMAP* approach including off-diagonal elements in the inverse covariance matrix (dashed red line); and a sharp transition between the low- and high-ℓ regions (dotted blue line).

(A color version of this figure is available in the online journal.)

is included twice, both by the low- and the high-ℓ likelihood, one must subtract the corresponding marginal for the transition region once to avoid double-counting (this is also an immediate consequence of Equation (1), under the assumption that $p(L|T, H) = p(L|T) = p(L, T)/P(T)$, i.e., the low-$\ell$ region is conditionally independent of the high-ℓ region given the transition region). Note that any estimator for the transition likelihood may be used for the correction term, typically by extracting the relevant range from either the low- or the high-ℓ likelihoods.

To assess the importance of the specific strategy adopted for hybridization, we modify the (7 yr) *WMAP* likelihood to include each of the three solutions, and derive constraints on the standard ΛCDM model using *WMAP* data only. The transition multipole is set to $\ell_{\rm trans} = 32$ for the sharp transition case, whereas the transition region is defined as $\ell = \{21, \ldots, 32\}$ for the new hybrid scheme. The *WMAP* BR estimator is used both for the low-ℓ and the transition regions in the latter case. We adopt $\Omega_b h^2, \Omega_m h^2, \theta, \tau, n_s,$ and $\log(10^{10} A_s)$ as our primary parameters, and adopt CosmoMC (Lewis & Bridle 2002) as our MCMC engine. The resulting one-dimensional marginals are shown in Figure 3 for all three cases, and posterior mean summary statistics are given in Table 1.

With a largest relative difference between any two cases of 0.06σ, these results demonstrate that the standard six-parameter

Table 1
Summary of Cosmological Parameters Derived with Three Different Hybridization Schemes

	Default *WMAP*	Sharp Transition		Transition Region	
	Constraint	Constraint	Deviation (σ)	Constraint	Deviation (σ)
$\Omega_b h^2$	0.0225 ± 0.0006	0.0225 ± 0.0006	0.02	0.0225 ± 0.0006	0.02
$\Omega_m h^2$	0.111 ± 0.005	0.111 ± 0.005	0.01	0.112 ± 0.006	0.05
θ	1.039 ± 0.003	1.039 ± 0.003	0.04	1.039 ± 0.003	0.05
τ	0.088 ± 0.015	0.088 ± 0.015	0.04	0.088 ± 0.015	0.05
n_s	0.969 ± 0.013	0.969 ± 0.014	0.03	0.968 ± 0.014	0.06
$\log[10^{10} A_s]$	3.08 ± 0.04	3.08 ± 0.03	0.03	3.08 ± 0.04	0.05

Notes. The three hybridization schemes are as follows. The original *WMAP* approach including off-diagonal elements in the inverse covariance matrix (second column), a sharp transformation at $\ell_{\rm trans} = 32$ (third column), and the new approach implementing a transition region between $\ell = 21$ and 32 (fifth column). The confidence intervals are 1 σ, and the best-fit points are the marginalized means of the parameters. The fourth and sixth columns show the relative shifts with respect to the *WMAP* approach measured in units of σ.

ΛCDM model is highly robust with respect to assumptions about the correlations across the transition regime. Similar conclusions were found when performing an identical analysis for the recently released *Planck* likelihood (Planck XV 2013), and this motivated the choice of a sharp transition for that particular implementation. For future experiments and analyses we nevertheless recommend the hybrid approach presented here, for two main reasons. First, our expression provides a statistically well motivated solution whose validity may be monitored directly through the C_ℓ covariance matrix; without the same level of statistical rigor, detailed simulations are more critical for the other two approaches, and these should in principle be repeated both when the data set or the parametric model is changed. Second, this expression is implementationally trivial once both low- and high-ℓ likelihoods are available, and there is therefore no practical reason for not including these correlations, even if their impact may be small.

4. FASTER BLACKWELL–RAO CONVERGENCE

4.1. Review of the Blackwell–Rao Estimator

As mentioned in Section 1, both the *Planck* and *WMAP* low-ℓ likelihoods (Planck XV 2013; Hinshaw et al. 2013) employs a specific BR estimator to produce an accurate likelihood approximation that accounts for all correlations and non-Gaussian structures (Chu et al. 2005). The main advantages of this estimator are (1) computational speed, (2) implementational simplicity, and (3) support for seamless marginalization over systematic effects and component separation errors through Gibbs sampling (Jewell et al. 2004; Wandelt et al. 2004; Eriksen et al. 2004).

This estimator may be explained intuitively as follows: Suppose it is possible to construct an experiment that provides a perfect full-sky noiseless image of the CMB sky, $\mathbf{d} = \mathbf{s}$. For that experiment, the only source of uncertainty on C_ℓ is cosmic variance, and the exact CMB likelihood in Equation (5) reduces to an inverse Gamma distribution,

$$\mathcal{L}_0(C_\ell) \propto \frac{e^{-\frac{1}{2}\mathbf{s}^t \mathbf{S}(C_\ell)^{-1}\mathbf{s}}}{\sqrt{|\mathbf{S}(C_\ell)|}} \propto \prod_\ell \sigma_\ell^{-\frac{2\ell-1}{2}} \frac{e^{\frac{2\ell+1}{2}\frac{\sigma_\ell}{C_\ell}}}{C_\ell^{\frac{2\ell+1}{2}}}. \quad (8)$$

Here we have defined $\sigma_\ell \equiv (1/(2\ell+1)) \sum_{\ell=-m}^{m} |a_{\ell m}|^2$ to be the realization specific power spectrum of $\mathbf{s}$.

However, for any real experiment there are additional sources of uncertainty beyond cosmic variance, for instance from instrumental noise and foreground contamination, and $P(\mathbf{s}|\mathbf{d})$ is no longer a delta function. In order to account for this additional uncertainty, one must weight the ideal likelihood in Equation (8) with respect to $P(\mathbf{s}|\mathbf{d})$,

$$\mathcal{L}_{\rm BR}(C_\ell) = \int d\mathbf{s}\, \mathcal{L}_0(C_\ell)\, P(\mathbf{s}|\mathbf{d}). \quad (9)$$

At first glance, this integral appears difficult to evaluate, as it involves millions of degrees of freedom. However, this is precisely where the CMB Gibbs sampler enters the picture. As explained in detail by Jewell et al. (2004), Wandelt et al. (2004), and Eriksen et al. (2004), the output from this algorithm is a set of samples drawn directly from $P(\mathbf{s}|\mathbf{d})$, accounting for both instrumental noise and foreground errors. Thus, the integral can be simply evaluated by Monte Carlo integration as a sum over these samples,

$$\mathcal{L}_{\rm BR}(C_\ell) \approx \sum_{i=1}^{N_{\rm samp}} \prod_{\ell=\ell_{\rm min}}^{\ell_{\rm max}} \sigma_\ell^{i\,\frac{2\ell-1}{2}} \frac{e^{\frac{2\ell+1}{2}\frac{\sigma_\ell^i}{C_\ell}}}{C_\ell^{\frac{2\ell+1}{2}}}. \quad (10)$$

This is the CMB power spectrum BR estimator, which is guaranteed to converge to the true likelihood in the limit of $N_{\rm samp} \to \infty$.

4.2. Lifting the "Curse of Dimensionality" by Block Factorization

While the BR estimator is guaranteed to converge to the correct answer, it is not obvious how fast it does so, as measured in terms of number of samples required for convergence, $N_{\rm samp}$. Further, since the computational cost of a single Gibbs sample is typically on the order of several CPU hours (Eriksen et al. 2004), depending on the angular resolution and/or signal-to-noise ratio of the data set under consideration, it is important to understand this scaling before attempting a full-scale analysis. Indeed, Chu et al. (2005) showed that $N_{\rm samp}$ scales exponentially with $\ell_{\rm max}$, effectively limiting its operational range to $\ell_{\rm max} \approx 50$–$70$. The main goal of the present section is to improve on this limit, and extend the BR estimator to high ℓ's.

To understand the origin of the exponential scaling, we show in Figure 4 a simple two-dimensional Gaussian distribution mapped by a Monte Carlo sampler. The top and left panels show the respective one-dimensional marginals. The BR estimator establishes a smooth approximation to these distribution by assigning a kernel of finite width to each individual Monte Carlo sample (illustrated by blue contours/Gaussians) before taking the average over all samples. Suppose now that the width of the one-dimensional kernel is 10% of the width of the marginal

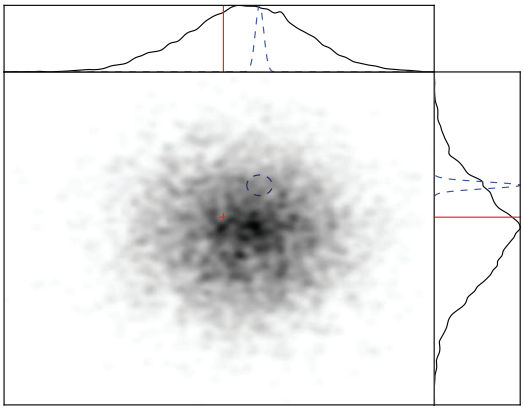


Figure 4. Illustration of the "curse of dimensionality." The Blackwell–Rao estimator builds up a smooth histogram from a finite set of Monte Carlo samples by assigning a distribution (or kernel) to each sample. The number of samples required to reach convergence is proportional to the ratio between the volume of the kernel (blue) and the volume of the full distribution (black). If this ratio is $r < 1$ in one dimension (top and left panels), it is r^2 in two dimensions (central panel), and r^n in n dimensions. This implies that the number of Monte Carlo samples required to reach convergence for the CMB BR estimator scales exponentially with $\ell_{\rm max}$. The evaluation of the two-dimensional likelihood at a specific point in parameter space (red cross) will be much more sensitive to the number of samples than the corresponding evaluations in the respective marginalized parameter spaces (red lines).

(A color version of this figure is available in the online journal.)

distribution; in that case, one needs $\sim$10 samples in order to cover the marginal once. In two dimensions, however, one needs $\sim 10^2$ samples to cover the full joint distribution once, since the ratio now is only 10% in each of the two directions. More generally, in n dimensions one would need $\sim 10^n$ samples. This is a variation of the well-known "curse of dimensionality," which says that the number of points required to cover an n-dimensional space scales exponentially with n.

The BR estimator given in Equation (10) converges well up to $\ell \approx 30$ with only a few thousand samples for *WMAP* (Chu et al. 2005), while for *Planck* it is found to be robust up to $\ell \approx 70$ with 100,000 samples (Planck XV 2013). To extend to even higher ℓ's by brute force would soon require a prohibitively large number of samples, as the computational cost for the Gibbs sampling step of the latter case is already half a million CPU hours.

Fortunately, the block factorization presented in Section 2 may be used to define an alternative and computationally much cheaper algorithm.

1. Partition the full $\ell_{\rm max}$-dimensional $\mathcal{L}(C_\ell)$ into a sequence of lower-dimensional blocks, r_k. Here, we take the blocks to be of the same width, which we call $\Delta\ell$.
2. Use the standard BR estimator to estimate the marginal likelihood for each block and each neighboring set of two blocks.
3. Merge these block marginals into a single all-ℓ estimator through the block factorization in Equation (1).

Thus, our new likelihood approximation can be written succinctly on the following form,

$$\mathcal{L}(C_\ell) \approx \frac{\prod_{k=1}^{n-1} \mathcal{L}_{\rm BR}(r_k, r_{k+1})}{\prod_{k=2}^{n-2} \mathcal{L}_{\rm BR}(r_k)}. \tag{11}$$

Note that all the likelihood evaluations on the right side of this expression involve a maximum of $2\Delta\ell - 1$ dimensions, as opposed to $\ell_{\rm max} - \ell_{\rm min} + 1$ for the full joint BR estimator, effectively lifting the curse of dimensionality.

4.3. Accuracy and Convergence

4.3.1. Methodology

Before the block factorized BR estimator can be used for real analysis, it is necessary to assess its accuracy and convergence properties. To this aim, we analyze two different simulations with the above machinery, adopting the convergence analysis methodology of Chu et al. (2005). We use this to perform various tests which will be reported in the "Results" section. Monte Carlo samples are produced with Commander (Eriksen et al. 2004, 2008).

The first simulation consists of a full-sky high-resolution ($N_{\rm side} = 512$, $\ell_{\rm max} = 1024$, $14'$ Gaussian beam) data set with uniform noise (65 μK rms per pixel). The main advantage of this case is that the C_ℓ likelihood (Equation (5)) factorizes in ℓ, and can be evaluated analytically,

$$\mathcal{L}_{\rm ideal}(C_\ell) \propto \prod_\ell \frac{e^{-\frac{2\ell+1}{2}\frac{\hat{\sigma}_\ell}{(C_\ell+N_\ell)}}}{(C_\ell + N_\ell)^{\frac{2\ell+1}{2}}}, \tag{12}$$

where $\hat{\sigma}_\ell$ is the angular power spectrum of the noisy sky map, and N_ℓ is the ensemble averaged noise power spectrum. The second simulation consists of a low-resolution ($N_{\rm side} = 32$, $\ell_{\rm max} = 95$, 6° FWHM Gaussian beam) data set with various sky masks imposed. White noise of 0.3 μK rms is added to each pixel, resulting in a signal-to-noise of unity at $\ell \approx 70$. The main purpose of this simulation is to study the effect of correlations between different multipoles arising from the sky cut through comparison with brute-force pixel-space likelihood evaluation. However, because of the brute-force evaluations, this case is necessarily limited to low angular resolution.

The CMB signal is drawn from a Gaussian distribution with a covariance given by the best-fit *WMAP* ΛCDM power spectrum, $C_\ell^{\rm ref}$ (Hinshaw et al. 2013). In each case, we fit a two-parameter amplitude-tilt (A–n) model on the form

$$C_\ell(A, n) = A \left(\frac{\ell}{\ell_0}\right)^n C_\ell^{\rm ref}, \tag{13}$$

where $\ell_0 = \ell_{\rm max}/2$, simply by mapping out $\mathcal{L}(A, n)$ over a two-dimensional grid. For $\ell_{\rm min} = 2$, this choice of pivot multipole ensures a low degree of correlation between A and n.

To assess both convergence and accuracy, we adopt the following measure of difference between two likelihoods, $\mathcal{L}_1$ and $\mathcal{L}_2$ (Chu et al. 2005),

$$q = \int |\mathcal{L}_1(A, n) - \mathcal{L}_2(A, n)|\, dA\, dn. \tag{14}$$

One can show that if $\mathcal{L}_1$ and $\mathcal{L}_2$ are two bivariate Gaussian distributions with the same covariance matrix, $\mathbf{\Sigma}$, but different means, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$, then

$$q = \Phi\left(\tfrac{1}{2\sqrt{2}}\sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\mathbf{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}\right), \tag{15}$$

where Φ is the cumulative standard normal distribution function. From this, one finds that a 0.1σ shift in a Gaussian distribution

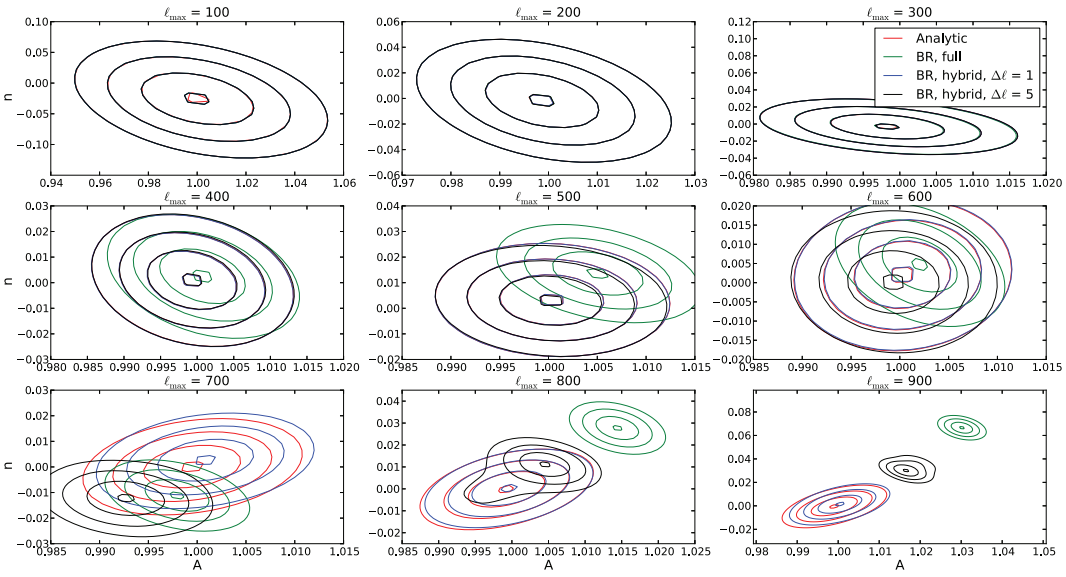


Figure 5. Comparison of four different methods of evaluating a simple amplitude-tilt likelihood for a full-sky simulation: The analytic case, the full Blackwell–Rao case, and two versions of the hybrid likelihood described in this paper—with $\Delta\ell = 1$ and 5, respectively.

corresponds to $q \sim 0.05$. In the following, we therefore define two distributions to agree if $q < 0.05$.

For the accuracy assessment, we simply compare the block factorized BR likelihood with the exact case. Convergence assessment, however, is done by drawing two disjoint sample subsets from the full set of available Monte Carlo samples, compute the BR estimator from each subset, and compare the resulting likelihoods. We then increase the number of samples in the two subsets, $N_{\rm samp}$, until q is consistently lower than 0.05 even when adding 100 additional samples; the latter criterion is imposed in order to avoid chance agreement. Finally, we repeat this calculation a certain number of times with different sample subsets (but drawn from the same full sample set), and report the median of the resulting values of $N_{\rm samp}$ as the final estimate of the number of samples required for convergence.

4.3.2. Results

Figure 5 shows $\mathcal{L}(A, n)$ evaluated from the high-resolution full-sky simulation for nine different values of $\ell_{\rm max}$ with four different likelihood expressions; analytic, standard BR, and two variations of the block-factorized BR estimator. A total of $N_{\rm samp} = 28{,}000$ samples are included in the two latter, a choice that is set to highlight the fundamental difference between the various cases. In particular, since there are no correlations between any multipoles in this case, all four approaches are in principle exact, and the only difference among the four cases are their relative convergence rates.

For $\ell_{\rm max} \leqslant 300$, we see that all four estimators agree to very high accuracy. However, from $\ell_{\rm max} \geqslant 400$ the full-range BR likelihood away starts to diverge from the others. At $\ell_{\rm max} = 900$, it is separated from the analytic result by more than 15σ. In this case, the sum in Equation (10) is strongly dominated by the one sample that happens to have the lowest power spectrum scatter about some best-fit mode, and the resulting distribution is simply an imprint of the cosmic variance kernel (Equation (8)) for that sample.

The block factorized BR estimators remain valid to higher $\ell_{\rm max}$, demonstrating how the "curse of dimensionality" is lifted by breaking the full parameter space into smaller regions that are easier to handle. In particular, the case with $\Delta\ell = 1$ agrees with the analytic case even at $\ell_{\rm max} = 900$ to $\sim 0.3\sigma$.

Next, in the top panel of Figure 6 we plot the number of samples required for convergence according to the above criterion for the high-resolution full-sky simulation described above, and in the bottom panel we show the same, but after applying the *WMAP* mask, covering 25% of the sky, in order to introduce a realistic multipole correlation structure. The upper vertical limit in these plots is set by the finite number of samples included in the analysis.

In all cases we see the same qualitative behavior. Reducing the dimensionality of the BR estimator through block factorization greatly improves the convergence rate by reducing the required number of samples by orders of magnitude at high ℓ's. For instance, for the full-sky case and with a block size of $\Delta\ell = 6$, only 10^3 samples are required in order to reach convergence up to $\ell_{\rm max} = 500$, whereas the full BR estimator would require 10^6. For the 25% *WMAP* mask, about 10^4 samples are required for $\ell_{\rm max} = 200$, while it is difficult to establish any sensible estimate for the full BR estimator in this case. (Note that the high-ℓ projection for the latter case, marked by a dashed line, is based on linear extrapolation from a few low-ℓ points, since convergence was not reached at all within the current sample set at higher multipoles. This projection is therefore associated with a very large systematic uncertainty.)

Finally, in Figure 7 we illustrate how the robustness of the split BR estimator depends on the block size chosen, for a fixed number of samples. We use a symmetric mask, shown in Figure 8, which covers a 20 deg strip of the galactic center. It is close in extent to the COBE mask. We then apply the

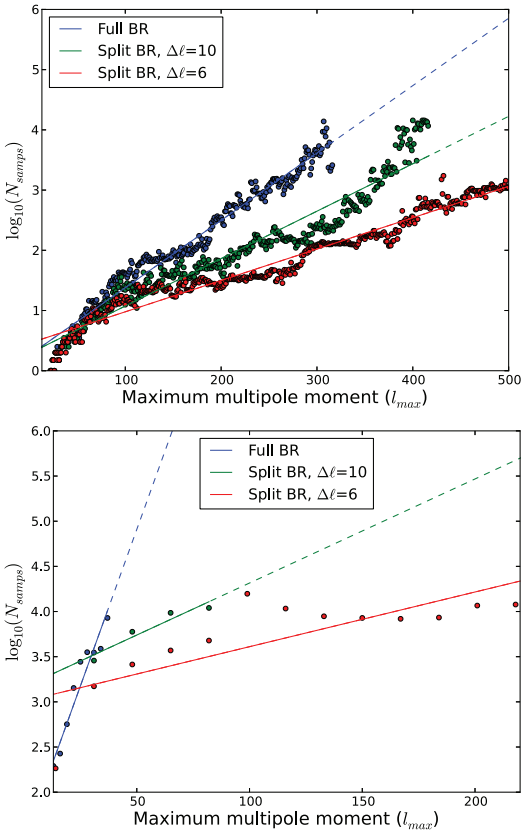


Figure 6. Convergence analysis for the split Blackwell–Rao estimator, with convergence defined in Section 4.3. The samples come from running Commander on a full-sky simulation. We show the median of the number of samples needed for convergence for a given $\ell_{\max}$, along with the best-fit regression line in $\log_{10}$-space. The median is computed from 10 (top) and 1024 (bottom) runs where the samples are scrambled between each run. The regression lines are dotted when they extend past the available data points. The high number of runs per data point for the bottom plot is also the reason for the more sparse sampling—each data point represented a very high computational cost, and so the number of data points were reduced.

BR estimator, using 30,000 Gibbs samples based on the low-resolution simulation. We also use the pixel-based estimator and calculate the q convergence between the BR result and the pixel-based result. Varying $\Delta\ell$, we the get a sense of how this block size affects the result. For this likelihood evaluation, we used $\ell_{\max} = 40$. A $\Delta\ell = 0$ on the plot means that the full BR estimator was used.

Conceptually, there should be two effects going on in this plot that both serve to reduce accuracy: When $\Delta\ell$ is too small, the between-multipole correlations set up by the mask are not modeled well enough, whereas when $\Delta\ell$ is too large, the parameter space becomes so large as to reduce accuracy. We can see this effect in play in the plot, although these data points are only indicative, not conclusive or exhaustive in any way. The specific block size dependence will typically depend on both the actual CMB signal and the morphology of the mask, and so prior testing like the methods outlined above should be performed before deciding on the optimal block size.

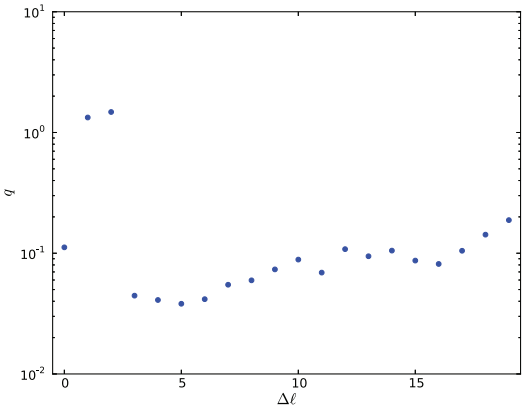


Figure 7. q-statistic (defined in the text) calculated for various modes of the Blackwell–Rao estimator compared with a pixel-based evaluation. The mask used is shown in Figure 8. A $\Delta\ell = 0$ means that the full BR estimator was used, while for $\Delta\ell > 0$ we have used the split BR estimator with a block size of $\Delta\ell$. The $\ell_{\max}$ used in the likelihood evaluation is 40.

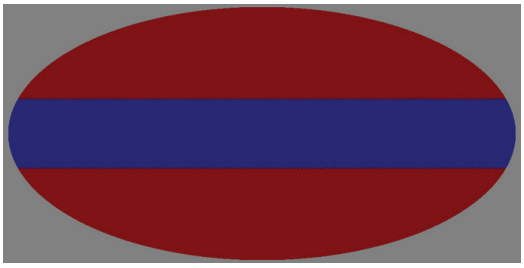


Figure 8. Mask used to test robustness in this section, covering a 20 deg strip centered on the galactic center.

We also tried applying the above method to larger masks, covering 30 deg and more, but the number of samples needed for convergence then quickly rose beyond 30,000. This shows that for masks that are significantly larger than those currently used in full-sky experiments, this likelihood estimator should not be trusted without prior testing—especially with regard to the number of Gibbs samples needed for a robust evaluation.

5. CONCLUSIONS

The main result presented in this paper is a statistically well motivated block factorization of the CMB power spectrum likelihood. Because the spherical harmonics are nearly orthogonal over the large sky coverages achieved by current CMB satellite experiments such as *Planck* and *WMAP*, any correlations between different C_ℓs are localized in multipole space. Under the assumption that these probabilistic dependencies have a strictly finite range, the full CMB likelihood may be reduced into a product of lower-dimensional marginals.

We have applied this result to two outstanding problems in CMB analysis. First, we use this expression to derive a well-motivated hybrid CMB likelihood estimator, merging an exact low-ℓ component with an approximate high-ℓ component, that accounts for correlations between the two regions. Although a detailed analysis of the *WMAP* likelihood shows that these

correlations are negligible for the *WMAP* sky cut and the six-parameter ΛCDM model, we nevertheless recommend this new estimator for future experiments and analyses, both because its implementation is trivial, and because it provides additional safety when analyzing non-standard models.

Second, we have shown how the same expression may be used to accelerate the convergence rate of the BR CMB likelihood estimator by orders of magnitude at high ℓs. This is achieved by factorizing the full parameter space into subspaces that each individually converge faster, and then merging these sub-blocks into a full-range estimator at the likelihood level using the block factorization formula.

It should be noted that these results rely directly on the assumption of vanishing long-range correlations. While this assumption holds to a very high accuracy for the basic CMB signal plus noise data model, it is in general not valid when including systematic effects in the analysis. Perhaps the two most important examples are correlated beam uncertainties and unresolved extra-Galactic point sources, each of which extend through all ℓ's (e.g., Planck XV 2013). Fortunately, these long-range degrees of freedom may be modeled in terms of a small number of power spectrum templates, each with an unknown amplitude. One can therefore marginalize over these by sampling the unknown amplitudes as nuisance parameters, similar to what was done for high-ℓ astrophysical parameters in the 2013 *Planck* likelihood (Planck XV 2013).

Finally, we note that the block factorization presented in Section 2 is a completely general statistical result that holds exactly for any banded probability distribution, and we therefore expect it to also find applications outside the CMB field.

This project was supported by the ERC Starting Grant StG2010-257080. Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. Some of the results in this paper have been derived using the HEALPix (Górski et al. 2005) software and analysis package.

REFERENCES

Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003a, ApJS, 148, 1
Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2003b, ApJS, 148, 97
Chu, M., Eriksen, H. K., Knox, L., et al. 2005, PhRvD, 71, 103002
Eriksen, H. K., Jewell, J. B., Dickinson, C., et al. 2008, ApJ, 676, 10
Eriksen, H. K., O'Dwyer, I. J., Jewell, J. B., et al. 2004, ApJS, 155, 227
Górski, K. M. 1994, ApJL, 430, L85
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hinshaw, G., Spergel, D. N., Verde, L., et al. 2003, ApJS, 148, 135
Hivon, E., Górski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2
Jewell, J., Levin, S., & Anderson, C. H. 2004, ApJ, 609, 1
Lewis, A., & Bridle, S. 2002, PhRvD, 66, 103511
Penzias, A. A., & Wilson, R. W. 1965, ApJ, 142, 419
Planck Collaboration I 2013, arXiv:1303.5062
Planck Collaboration XII 2013, arXiv:1303.5072
Planck Collaboration XV 2013, 1303.5075
Planck Collaboration XXIII 2013, arXiv:1303.5083
Rocha, G., Contaldi, C. R., Bond, J. R., & Górski, K. M. 2011, MNRAS, 414, 823
Rudjord, Ø., Groeneboom, N. E., Eriksen, H. K., et al. 2009, ApJ, 692, 1669
Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1992, ApJL, 396, L1
Verde, L., Peiris, H. V., Spergel, D. N., et al. 2003, ApJS, 148, 195
Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. 2004, PhRvD, 70, 083511

Paper V

Grid-based exploration of cosmological parameter space with Snake
Mikkelsen, K. and Næss, S. K. and Eriksen, H. K.
Astrophysical Journal, Volume 777, issue 2, article id. 172 (2013)

The Astrophysical Journal, 777:172 (6pp), 2013 November 10 doi:10.1088/0004-637X/777/2/172

GRID-BASED EXPLORATION OF COSMOLOGICAL PARAMETER SPACE WITH SNAKE

K. Mikkelsen[1], S. K. Næss[1], and H. K. Eriksen[1,2]
[1] Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029, Blindern, NO-0315 Oslo, Norway; kristin.mikkelsen@astro.uio.no
[2] Centre of Mathematics for Applications, University of Oslo, P.O. Box 1053, Blindern, NO-0316 Oslo, Norway
Received 2012 November 13; accepted 2013 September 23; published 2013 October 24

ABSTRACT

We present a fully parallelized grid-based parameter estimation algorithm for investigating multidimensional likelihoods called Snake, and apply it to cosmological parameter estimation. The basic idea is to map out the likelihood grid-cell by grid-cell according to decreasing likelihood, and stop when a certain threshold has been reached. This approach improves vastly on the "curse of dimensionality" problem plaguing standard grid-based parameter estimation simply by disregarding grid cells with negligible likelihood. The main advantages of this method compared to standard Metropolis–Hastings Markov Chain Monte Carlo methods include (1) trivial extraction of arbitrary conditional distributions; (2) direct access to Bayesian evidences; (3) better sampling of the tails of the distribution; and (4) nearly perfect parallelization scaling. The main disadvantage is, as in the case of brute-force grid-based evaluation, a dependency on the number of parameters, $N_{\rm par}$. One of the main goals of the present paper is to determine how large $N_{\rm par}$ can be, while still maintaining reasonable computational efficiency; we find that $N_{\rm par} = 12$ is well within the capabilities of the method. The performance of the code is tested by comparing cosmological parameters estimated using Snake and the *WMAP*-7 data with those obtained using CosmoMC, the current standard code in the field. We find fully consistent results, with similar computational expenses, but shorter wall time due to the perfect parallelization scheme.

Key words: cosmic background radiation – cosmology: observations – methods: statistical

Online-only material: color figures

1. INTRODUCTION

Cosmological models are described in terms of a modest number of cosmological parameters that reflect the underlying physical processes of the universe. These are today routinely measured by experiments such as the *Wilkinson Microwave Anisotropy Probe* (*WMAP*; Jarosik et al. 2011), Planck Collaboration (2011) and the Sloan Digital Sky Survey (York et al. 2000) through likelihood techniques.

The most popular parameter estimation algorithm in the cosmology community to date is the CosmoMC package (Lewis & Bridle 2002), which maps out the cosmological parameter space using a Metropolis–Hastings Markov Chain Monte Carlo (MCMC) sampler. The computational cost of this method is almost exclusively determined by the external evaluation of the likelihood, which typically takes a few seconds per evaluation; the expense of the internal book-keeping operation is completely negligible compared to this. A complete analysis of current data sets typically requires $\mathcal{O}(10^5)$ evaluations, resulting in an overall computational cost of 100–10,000 CPU hours, depending on the particular problem.

This process can be sped up in two fundamentally different ways, namely either by reducing the cost per likelihood evaluation, or by reducing the number of likelihood evaluations required, and both cases have already been explored extensively in the literature. Examples of the former include CMBFit (Sandvik et al. 2004), PICO (Fendt & Wandelt 2007), COSMONET (Auld et al. 2007), sparse grids (Frommert et al. 2010), and PkANN (Agarwal et al. 2012), all of which essentially build up a library of known cosmological models given a set of parameters, and interpolate within this library using some statistical method. Examples of the latter include MultiNest (Feroz et al. 2009), APS (Daniel et al. 2012), CosmoPMC (Kilbinger et al. 2011), and HMC (Hajian 2007), all of which reduce the number of likelihood evaluations through more efficient sampling algorithms than the Metropolis–Hastings sampler.

In this paper, we present an algorithm that falls in the last category, aiming to reduce the total number of likelihood evaluations rather than the cost per evaluation. The initial idea of this paper is based on the following reasoning: if the problem under consideration involved only a one-dimensional likelihood, the mapping algorithm of choice would be obvious—one would simply evaluate the likelihood over a one-dimensional grid. The resulting function is both easier to work with than a set of samples, as produced by an MCMC algorithm, and more accurate. Furthermore, it generally requires fewer evaluations, because whereas an MCMC approach builds up the shape of the distribution by counting how many samples fall in a given parameter range ("bin"), the direct approach only needs to evaluate the likelihood in a given bin once. In other words, the MCMC approach spends most of the time evaluating the same likelihood points over and over again, which can give the direct evaluation approach a computational edge.

The vast majority of two-dimensional likelihoods are also mapped by grid methods rather than MCMC methods, while for three or four dimensions, the preferred approach is not clear. However, for higher dimensions, virtually all cases are so far handled by MCMC methods. At this stage, the so-called curse of dimensionality becomes highly relevant, as the number of likelihood evaluations depends exponentially on the number of dimensions. For instance, computing 100 grid points in each of five dimensions requires 100^5 evaluation, which is generally far too many for most problems.

However, in this paper we point out that this is not necessarily true. The point is simply that the vast majority of the high-dimensionality volume typically has negligible likelihood, and therefore does not need to be evaluated in the first place. The trick is to figure out which grid cells are relevant and which

are not. If this can be done both efficiently and robustly, all the useful properties of normal grids are retained, and computational cost is not compromised. Further, by virtue of not being a Markov chain, the algorithm parallelizes trivially, leading to shorter overall computational wall time, which is often even more critical for a given analysis problem than the total CPU time.

2. THE SNAKE ALGORITHM

2.1. Algorithm

The Snake algorithm is very simple, and can easily be explained in terms of a few basic steps. To do so succinctly, it is useful to first define some terminology.

The grid. The Snake algorithm operates on a virtual grid in parameter space with the origin, θ_0, set to the starting point selected by the user. This point together with the cell size in each dimension, $\mathbf{\Delta}\theta$, defines the position of every other grid point in parameter space through $\theta_0 + \mathbf{k} \cdot \mathbf{\Delta}\theta$, where $\mathbf{k}$ is an integer vector describing the cell location with respect to the origin.

The surface. Each point on the grid is assigned to one of three groups depending on whether they are external, internal, or surface points. External points are those that have not yet been considered; internal points are those for which the point itself and *all* its neighbors have been visited; and surface points are the points that have been considered, but have at least one unexplored neighbor.

The repository. Considered parameter points are assigned a book-keeping integer, i, representing that this point was the ith point visited by Snake. These points are stored as objects in a data structure called a repository, which is a two-dimensional dynamic list in which each row defines a point on the grid. In addition to the integer i each data row contains the point's $\mathbf{k}$ vector, likelihood value $\mathcal{L}_i$, the integers corresponding to its neighbors, and a logical flag specifying whether the point is currently on the surface.

Given these definitions, the Snake algorithm may be summarized as follows.

1. *Initialization.* Compute the likelihood of the starting point, and insert the required information into the repository.
2. *Neighbor.* Consider the surface point with the highest likelihood value, which has index i_S, and randomly pick one of its unexplored neighbors. Evaluate the likelihood of this new point.
3. *Surface update.* Find the integers corresponding to all the neighbors of the new point. If any of these neighbors or the new point no longer have any unexplored neighbors, set its surface flag to false and remove it from the list of surface points.
4. *Update repository.* Insert the new point into the repository structure and update the relevant information for its neighbors.
5. *Convergence check.* The grid point with the overall highest likelihood, with $i = i_{\rm peak}$, and the current surface points are used to test if Snake has converged. If $\log \mathcal{L}(i_{\rm peak}) - \log \mathcal{L}(\theta(i))$ is smaller than a predefined threshold for all surface points Snake has converged and the routine exits. Otherwise Snake loops back to (2) and performs another evaluation.

This stepping procedure leads to two distinct phases. First there is a burn-in period in which Snake performs a greedy maximum-likelihood search to find the peak of the likelihood distribution with index $i_{\rm peak}$. Then, once the maximum has been located, the area around the peak is investigated by stepping to an unexplored neighbor of the surface point with the highest likelihood. This ensures that the surface grows outward according to the underlying likelihood distribution, and continues until all the surface points have reached the threshold. A large threshold value lets Snake investigate a larger parameter volume which ensures that the tails are investigated more closely; however, this requires more evaluations to be performed which is computationally expensive. The threshold should therefore be kept low enough to keep cost reasonable, but still large enough to make sure the edges are properly investigated.

The likelihood evaluation is by far the most time consuming component of cosmological evaluations, which implies that efficiency can be increased by parallelizing the Snake algorithm. In the current implementation a master–slave parallelization strategy has been adopted, in which one processor maintains the repository, and the remaining processors only perform likelihood evaluations for parameters provided by the master. This ensures both a simple implementation as well as close to perfect speed-up; after only a few initial iterations there are always more than enough available surface points to keep all processors occupied. Moreover, the communication between the master and slaves is minimal, consisting only of a parameter multiplet and a likelihood return value.

As should be clear from the above, Snake is algorithmically trivial; it is nothing but an old-fashioned grid evaluation with efficient stepping and convergence. The only somewhat intricate part is to implement adequate book-keeping, which is necessary in order to maintain computational efficiency as the number of data elements, V, in the repository increases. For this purpose, we implement dictionaries, based on the C++ standard map template. These maps store the combination of two values, the key and the mapped value, and enable access to the mapped value by using the corresponding key in constant time, as opposed to $\mathcal{O}(V)$ for unsorted lists or $\mathcal{O}(\log V)$ for sorted lists.

Two such maps are implemented in Snake. The first for keeping track of which point in parameter space corresponds to which iteration, which is used to check if the neighbors of the current point have already been visited, and if so returns their iteration index. The second map keeps track of the likelihood value corresponding to each iteration and is sorted according to descending likelihood such that the first point on the list will always be that with the highest likelihood. Therefore, the i_S index is simply the mapped value at the top of the map. When a point becomes an interior point the corresponding entries are removed from the two map structures in order to keep these as short as possible and to avoid getting stuck at the overall maximum likelihood.

2.2. Walk-through of a Two-dimensional Example

Before testing the algorithm on realistic cases, it is useful to walk through it step-by-step for a simple case, to gain some intuition for its behavior. In this section, we therefore first consider the small two-dimensional example illustrated in Figure 1 and Table 1. The unknown distribution to be mapped is marked in Figure 1 by dashed lines, corresponding to 1σ, 2σ, and 3σ contours, and the threshold to be reached is defined as the 3σ contour.

First, we initialize the code at $(0, 0)$, which in this case happened to lie slightly below and to the left of the

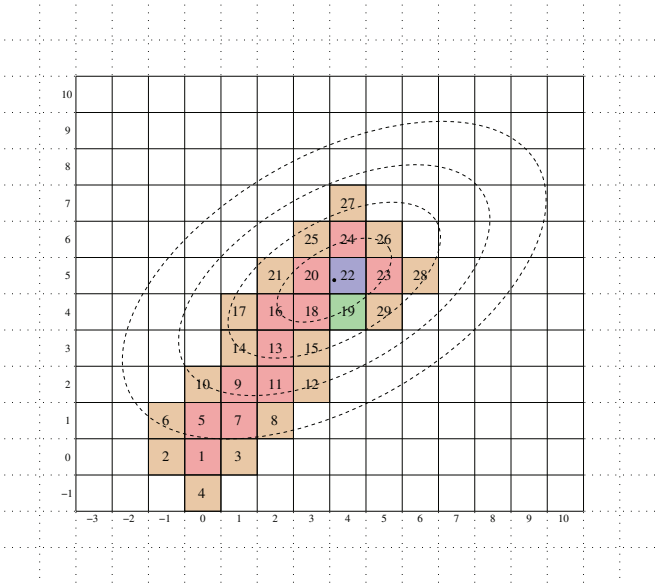


Figure 1. Snapshot of a made up likelihood distribution after 29 iterations illustrating the repository of Snake. Beige boxes are surface points, red are internal points, blue is the overall maximum likelihood, and green is the maximum likelihood on the surface.
(A color version of this figure is available in the online journal.)

Table 1
Two-dimensional Book Repository

i	**k**		$\mathcal{L}$	Ind				Surface
	x	y		$x:-1$	$x:+1$	$y:-1$	$y:+1$	
1	0	0	0.005	2	3	4	5	F
2	−1	0	0.003		1		6	T
3	1	0	0.002	1			7	T
4	0	−1	0.001				1	T
5	0	1	0.10	6	7	1	10	F
6	−1	1	0.05		5	2		T
7	1	1	0.15	5	8	3	9	F
8	2	1	0.10	7			11	T
22	4	5	0.95	20	23	19	24	F
23	5	5	0.75	22	28	29	26	F
24	4	6	0.70	25	26	22	27	F
25	3	6	0.60		24	20		T
26	5	6	0.70	24		23		T
27	4	7	0.45			24		T
28	6	5	0.55	23				T
29	5	4	0.55	19			23	T

maximum-likelihood point. We evaluate the likelihood, and insert this point into the first row of the repository (Table 1). At this stage, the first four columns are finalized, the surface flag is set to `true`, and none of the neighbor indices (indicated by the `ind` array of length $2N_{\rm par}$) are set, indicating that no neighbors have been evaluated yet.

Second, as specified by the algorithm, we now find the surface point with the highest likelihood, which of course is the point just added. We select one of its neighbors, which in this case happened to be $(-1, 0)$. We evaluate its likelihood, and insert this new point into the second row of the repository. We update the neighbor indices of both this new point and the original point to point to each others main index. We then repeat this process over and over again, adding more and more points to the repository, until the smallest difference between the likelihood of the overall maximum-likelihood point and that of any point on the surface is larger than a predefined threshold.

Table 1 gives a snapshot of the repository (parameters, likelihood, current status of the `ind` array, and the surface flag) at iteration number 29, matching the illustration seen in Figure 1. The beige boxes correspond to the points in parameter space which lie on the surface, red boxes are interior points, and the blue box corresponds to the overall maximum likelihood. The green box is the parameter point on the surface with the highest likelihood and will be the start point for the next iteration. The numbers inside the boxes correspond to the iteration index, thus the path Snake takes to reach the maximum likelihood can be seen, as well as the relation between neighbors and the values of the first and last eight points quoted in the `ind` array in Table 1. Iterations which have all `ind` columns filled have their surface flag set to false and the point no longer exists in the maps. The process continues until all boxes touching the 3σ contour have turned red, after which the surface lies fully below the threshold.

2.3. Exploration of Double-peaked Likelihood

A second illustration of how Snake investigates parameter space is given by the double-peaked two-dimensional likelihood

$$\mathcal{L} = A_1 e^{\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_1})^T C_1^{-1}(\boldsymbol{x}-\boldsymbol{\mu_1})} + A_2 e^{\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_2})^T C_2^{-1}(\boldsymbol{x}-\boldsymbol{\mu_2})}, \quad (1)$$

where $\boldsymbol{x}$ is the two-dimensional parameter vector, A_1 and A_2 are the peak amplitudes, C_1 and C_2 the corresponding covariance matrices, and $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$ the vectors of the means.

The leftmost plot of Figure 2 shows a likelihood distribution that can be described by this equation for a particular set of

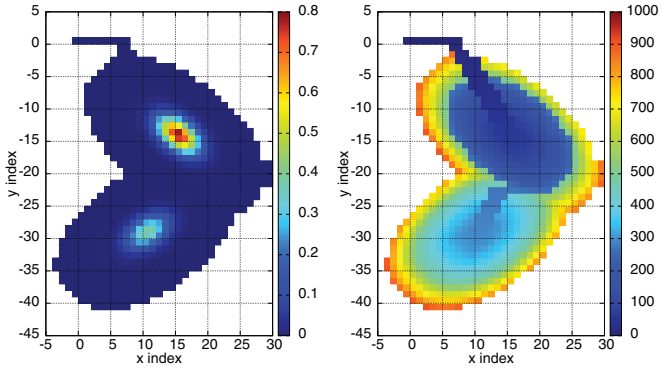


Figure 2. Two-dimensional illustration of Snake's sampling method. Left: the (unnormalized) target likelihood. Right: the path Snake takes through parameter space. It finds the closest peak, investigates the area around this peak, discovers the second peak, investigates the area around this one, and finally explores the joint boundary of both peaks.

(A color version of this figure is available in the online journal.)

covariance matrices and means. The path Snake takes in the two-dimensional parameter space is shown in the rightmost plot of Figure 2 and as can be seen, Snake quickly finds the maximum likelihood of the closest peak, and then proceeds by investigating the area around this peak by visiting neighbors of the surface point with highest likelihood. When the likelihood being investigated falls to the value corresponding to the intersection of the two peaks Snake makes its way to the second peak, and continues by investigating the area around this peak in the same manner as the first peak. Once Snake returns to the likelihood equal to that at the intersection it will investigate the points around both peaks until the desired threshold is reached.

Note that if the two peaks had been so far apart that the likelihood at the intersection fell below the threshold cutoff, the second peak would remain undiscovered. This problem can be solved in the same way as for standard Metropolis–Hasting samplers: run several Snakes in parallel with different initial positions. Once two independent Snakes touch for the first time, merge the repositories and the CPU working groups into one master–slave organization.

3. ACCURACY AND EFFICIENCY WITH INCREASING DIMENSIONALITY

The main outstanding question regarding the Snake algorithm is how well it scales with the number of dimensions in terms of efficiency. To study this question quantitatively, we consider a correlated Gaussian likelihood of the form

$$\mathcal{L} = e^{\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T C^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}, \quad (2)$$

where $\boldsymbol{x}$, C, and $\boldsymbol{\mu}$ are the multidimensional parameter vector, covariance matrix, and vector of means, respectively. For simplicity both the mean and standard deviation for dimension number i are chosen to equal the value i, thus $(\mu_i, \sigma_i) = (i, i)$.

Our goal is now to map out this distribution in $N_{\rm par}$ dimensions, and determine the maximum number of dimensions that can be probed with high accuracy using reasonable computational resources. To do so, we impose a limit on the number of likelihood evaluations of $N = 10^6$, a typical number for modern cosmological analyses. The grid cell width in dimension i is chosen to be $8i \times N^{1/N_{\rm par}}$, corresponding to distributing the N evaluations roughly over a grid covering roughly -4σ to $+4\sigma$ in each of the $N_{\rm par}$ dimensions. (Of course, the actual shape probed by Snake will not be a rectangular grid, but rather conform to the shape of the underlying distribution.) We then run the algorithm for increasing $N_{\rm par}$, and compare the resulting marginals to the known analytic input marginals; once the combined error in the derived mean or standard deviation is larger than 0.1σ, we consider the algorithm to have broken down as a result of the sparse sampling of the underlying distribution.

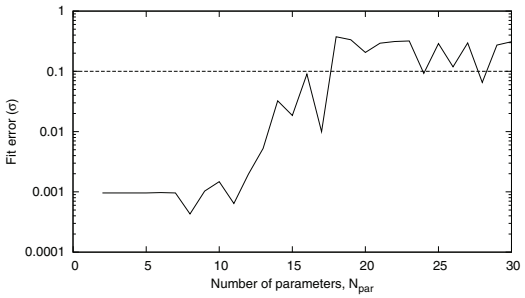


Figure 3. Combined mean and standard deviation errors averaged over number of dimensions (solid) showing the 0.1σ cutoff (dashed).

In Figure 3 we plot the combined mean and standard deviation errors averaged over the number of dimensions, $N_{\rm par}$, as a function of $N_{\rm par}$. Here we clearly see that for $N_{\rm par} < 12$, the algorithm recovers the true distribution with high accuracy. Of course, given more computational resources these errors can be decreased arbitrarily, but since the cost faces an exponential growth with increasing $N_{\rm par}$, it seems reasonable to define the operational range for Snake to be $N_{\rm par} \leqslant 12$–15.

4. SEVEN-YEAR *WMAP* LIKELIHOOD ANALYSIS

4.1. Parameter Estimation

We now apply this method to the seven-year *WMAP* likelihood, and estimate cosmological parameters within the well-established six-parameter ΛCDM concordance model (Komatsu et al. 2011). The parameter set of choice is $\Omega_b h^2$, $\Omega_{\rm DM} h^2$, θ,

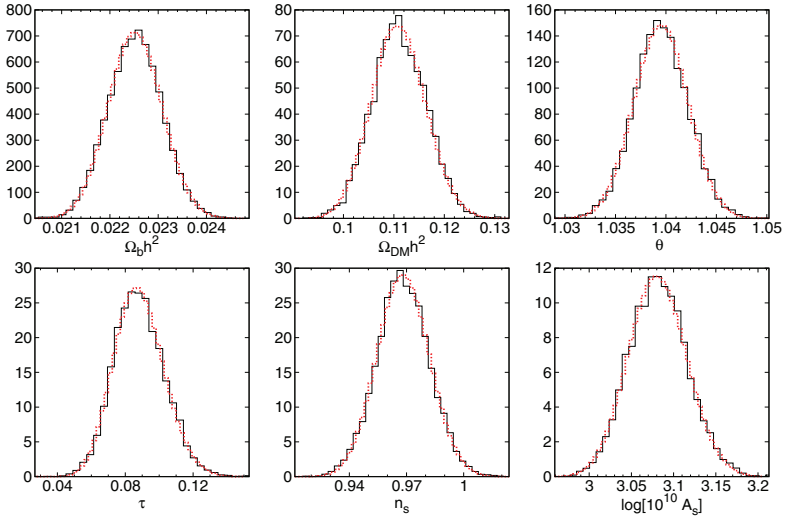


Figure 4. Marginal cosmological parameter distributions derived with Snake (dashed red line) and CosmoMC (solid black line) from the seven-year *WMAP* likelihood. (A color version of this figure is available in the online journal.)

Table 2
Cosmological Parameters

Parameter	CosmoMC	Snake	Shift in σ
$\Omega_b h^2$	$0.02252^{+0.00055}_{-0.00056}$	$0.02252^{+0.00057}_{-0.00056}$	0
$\Omega_{\rm DM} h^2$	$0.1110^{+0.0055}_{-0.0054}$	$0.1107^{+0.0055}_{-0.0054}$	0.06
θ	1.039±0.003	1.039±0.003	0
τ	$0.08849^{+0.00632}_{-0.00754}$	$0.08758^{+0.01558}_{-0.01426}$	0.08
n_s	$0.9682^{+0.0138}_{-0.0136}$	$0.9681^{+0.0139}_{-0.0138}$	0.07
log[$10^{10} A_s$]	$3.082^{+0.034}_{-0.035}$	3.080± 0.035	0.06

Notes. Comparison of best-fit parameters derived by CosmoMC and Snake from the seven-year *WMAP* data.

τ, n_s, and log[$10^{10} A_s$]. The same setup is analyzed using both Snake and CosmoMC for comparison purposes.

The resulting normalized marginal distributions are shown in Figure 4, and means and standard deviations are tabulated in Table 2. The agreement between the two methods is excellent, with a maximum difference between the two methods corresponding to a 0.08σ shift in τ and 0.07σ shift in n_s.

The CosmoMC results were obtained with an MPI convergence criterion of 0.03, while the Snake convergence threshold was defined to be -6.0. Both codes were run on 50 CPUs, and the resulting wall times were 1.42 and 1.24 hr, respectively.

4.2. Model Selection by Bayesian Evidence

A significant advantage of Snake over CosmoMC is its direct access to the Bayesian evidence (e.g., Gelman et al. 2003). For a given model H with parameters θ and data d, this is simply the normalization factor, $E \equiv P(d|H)$, in Bayes' theorem,

$$P(\theta|d, H) = \frac{P(d|\theta, H)P(\theta|H)}{P(d|H)}. \quad (3)$$

The other factors are the likelihood, $\mathcal{L}(\theta|H) = P(d|\theta, H)$, the prior, $P(\theta|H)$, and the posterior, $P(\theta|d, H)$. Different models can be compared in terms of their evidence, which for a model, H_n, is given by

$$P(d|H_n) = \int_\Omega P(d, \theta|H_n)\, d\theta = \int_\Omega P(d|\theta, H_n) P(\theta|H_n)\, d\theta, \quad (4)$$

where $P(d, \theta|H_n)$ is the joint probability distribution of d and θ given this model over all of parameter space, Ω, with step sizes of $d\theta$.

Calculating the evidence for different models using results from Snake is rather straightforward as the parameter space is gridded into even cells of volume $\int d\theta$. The integral in Equation (4) becomes a sum of the likelihood values within the threshold multiplied by the volume of one grid cell, where we assume a uniform prior which gives a factor of $1/L$ for each parameter, where L is the range for each parameter.

To compare two different models, H_1 and H_2, it is common to consider the quantity

$$\delta \log E = \log E_1 - \log E_2, \quad (5)$$

where E_1 and E_2 are the evidences of models H_1 and H_2, respectively. The larger the value of $\delta \log E$ the higher the evidence in favor of model E_1. To calibrate this quantity, one commonly adopts the Jeffreys' scale (Liddle et al. 2006; Trotta 2008),

$$\delta \log E > \begin{cases} 1 & \text{evidence for } E_1 \text{ is substantial} \\ 2.5 & \text{evidence for } E_1 \text{ is strong} \\ 5 & \text{evidence for } E_1 \text{ is decisive.} \end{cases}$$

However, one should note that this scale only provides a general guideline, and conclusions can be application specific; see, e.g., Nesseris & Garcia-Bellido (2013) for a recent discussion of this issue.

We now evaluate the evidence for both the standard six-parameter model described above and for the reduced model obtained by enforcing $n_s = 1$. We find that the individual evidences are $E_1 = -3743.21$ and $E_2 = -3744.62$, respectively,

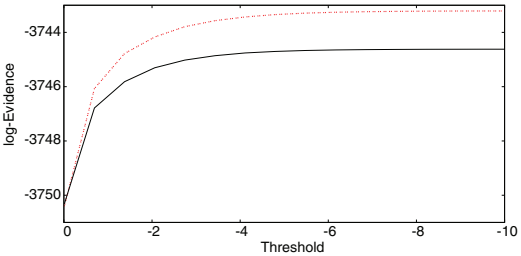


Figure 5. Evidence as a function of threshold for the standard six-parameter model (dashed red line) and for the five-parameter model with $n_s = 1$ (solid black line).

(A color version of this figure is available in the online journal.)

with an estimated uncertainty in each of 0.1. This corresponds to $\Delta \log E$ of 1.41 in favor of the six-parameter model; the full model therefore provides a better fit to the data, even when accounting for the larger parameter volume. Similar results have already been published by Parkinson & Liddle (2010).

As mentioned the evidence is dependent on the volume of the grid points investigated, which in turn is highly dependent on the preselected threshold. If the threshold is too low the volume covered by the analysis will be too small to compute the full evidence since the points excluded would have values high enough to contribute significantly to the evidence. In other words, one must select a sufficiently large threshold such that the values of excluded points would be insignificant, to ensure that the evidence has converged. Figure 5 shows the logarithm of the evidence as a function of threshold for the standard six-parameter model and for the reduced five-parameter model where we have imposed $n_s = 1$. As can be seen the log-evidence for the reduced model is fully converged at a threshold of -6.0 whereas that for the full model is very close to converged. Thus, a threshold value of -6.0 is large enough to ensure evidence convergence while still being low enough to keep the number of grid points as low as possible.

5. SUMMARY AND OUTLOOK

In this paper we have described a simple grid-based estimator for multi-dimensional likelihoods. This algorithm exploits the fact that by far most of the $N_{\rm par}$-dimensional parameter volume in a general likelihood has negligible contributions, and spends its computational resources only where the likelihood itself is significant. However, in contrast to standard MCMC methods, it only considers each parameter point once, relying on the actual value of the likelihood.

The main advantages of this method are (1) trivial extraction of arbitrary conditional distributions; (2) direct access to Bayesian evidences; (3) better sampling of the tails of the distribution; and (4) nearly perfect parallelization scaling. The main disadvantage is a computational cost increasing exponentially with $N_{\rm par}$. However, we have shown that the algorithm is fully capable of probing at least $N_{\rm par} \lesssim 12$–15 with reasonable computational resources, which is sufficient for current cosmological models.

In the current implementation the total cost of the method is comparable to that of CosmoMC for similar convergence criteria. However, the cost for a full Snake analysis can be vastly reduced by introducing adaptive grids, in which the grid cell depends on the local properties of the likelihood, such that high-significance regions are sampled more densely than the tail regions. The results from this extension will be reported in a future publication.

The computations presented in this paper were carried out on Titan, a cluster owned and maintained by the University of Oslo and NOTUR. H.K.E. acknowledges support from the ERC Starting Grant StG2010-257080.

REFERENCES

Agarwal, S., Abdalla, F. B., Feldman, H. A., Lahav, O., & Thomas, S. A. 2012, MNRAS, 424, 1409
Auld, T., Bridges, M., Hobson, M. P., & Gull, S. F. 2007, MNRAS, 376, L11
Daniel, S. F., Connolly, A. J., & Schneider, J. 2012, arXiv:1205.2708
Fendt, W. A., & Wandelt, B. D. 2007, ApJ, 654, 2
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Frommert, M., Pflüger, D., Riller, T., et al. 2010, MNRAS, 406, 1177
Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2003, Bayesian Data Analysis (2nd ed.; Chapman & Hall)
Hajian, A. 2007, PhRvD, 75, 083525
Jarosik, N., Bennett, C. L., Dunkley, J., et al. 2011, ApJS, 192, 14
Kilbinger, M., Benabed, K., Cappe, O., et al. 2011, arXiv:1101.0950
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Lewis, A., & Bridle, S. 2002, PhRvD, 66, 103511
Liddle, A., Mukherjee, P., & Parkinson, D. 2006, A&G, 47, 30
Nesseris, S., & Garcia-Bellido, J. 2013, JCAP, 08, 036
Parkinson, D., & Liddle, A. R. 2010, PhRvD, 82, 103533
Planck Collaboration 2011, A&A, 536, 1
Sandvik, H. B., Tegmark, M., Wang, X., & Zaldarriaga, M. 2004, PhRvD, 69, 063005
Trotta, R. 2008, ConPh, 49, 71
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579